GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:36:25 ; Search time 94.14 Seconds
(without alignments)
14.159 Million cell updates/sec

Title: US-09-367-714A-23
Perfect score: 52
Sequence: 1 KLLLKLLLKLLK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	12	18	AAW35149	Leu/Lys diastereom
2	52	100.0	12	18	AAW35152	Leu/Lys diastereom
3	52	100.0	12	19	AAW82847	Antipathogenic pep
4	52	100.0	12	19	AAW82850	Antipathogenic pep
5	52	100.0	12	19	AAW82856	Antipathogenic pep
6	52	100.0	12	21	AAB17413	Antipathogenic pep
7	52	100.0	12	21	AAB17416	Antipathogenic pep
8	52	100.0	12	21	AAB17483	Antipathogenic pep
9	52	100.0	12	21	AAB17485	Antipathogenic pep
10	52	100.0	13	18	AAW35231	Diastereomer pepti
11	52	100.0	13	21	AAB17482	Antipathogenic pep
12	52	100.0	14	19	AAW82854	Antipathogenic pep
13	52	100.0	37	19	AAW77378	Lytic peptide with
14	52	100.0	77	19	AAW82858	Antipathogenic pep
15	52	100.0	77	19	AAW82859	Antipathogenic pep
16	46	88.5	12	18	AAW35150	Leu/Lys diastereom
17	46	88.5	12	18	AAW35153	Leu/Lys diastereom
18	46	88.5	12	18	AAW35169	Leu/Lys diastereom
19	46	88.5	12	18	AAW35170	Leu/Lys diastereom
20	46	88.5	12	18	AAW35171	Leu/Lys diastereom
21	46	88.5	12	19	AAW82848	Antipathogenic pep
22	46	88.5	12	19	AAW82851	Antipathogenic pep
23	46	88.5	12	19	AAW82857	Antipathogenic pep
24	46	88.5	12	19	AAW82885	Antipathogenic pep
25	46	88.5	12	19	AAW82886	Antipathogenic pep
26	46	88.5	12	19	AAW82887	Antipathogenic pep
27	46	88.5	12	21	AAB17414	Antipathogenic pep
28	46	88.5	12	21	AAB17417	Antipathogenic pep
29	46	88.5	12	21	AAB17420	Antipathogenic pep
30	46	88.5	12	21	AAB17421	Antipathogenic pep
31	46	88.5	12	21	AAB17422	Antipathogenic pep
32	46	88.5	13	18	AAW35232	Diastereomer pepti
33	46	88.5	13	21	AAB17484	Antipathogenic pep
34	46	88.5	14	19	AAW82855	Antipathogenic pep
35	46	88.5	21	22	AAU03187	Membrane active sy
36	46	88.5	21	22	AAB60066	KL3 membrane activ
37	46	88.5	153	20	AAY29393	Sperm whale myoglo
38	43	82.7	73	21	AAB52057	Human secreted pro
39	41.5	79.8	15	19	AAW77384	Lytic peptide with
40	41	78.8	28	10	AAP91335	Amino acid sequenc
41	40	76.9	14	16	AAR67795	Antimicrobial olig
42	40	76.9	14	21	AAB17122	Calmodulin antagon
43	40	76.9	15	15	AAR56957	Peptide which neut
44	40	76.9	16	16	AAR67797	Bisamine derivatiz
45	40	76.9	16	16	AAR67798	Monoamine derivati

ALIGNMENTS

RESULT 1
AAW35149
ID AAW35149 standard; peptide; 12 AA.
XX
AC AAW35149;
XX
DT 14-APR-1998 (first entry)
XX
DE Leu/Lys diastereomer peptide [D]-L3,4,8,10-K4L8.
XX
KW Leu/Lys diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW Honey bee venom; pardaxin; cytolytic activity; cancer;
KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW agricultural pesticide; cell wall lysis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 3
FT /note= "D-form residue"
FT Misc-difference 4
FT /note= "D-form residue"
FT Misc-difference 8
FT /note= "D-form residue"
FT Misc-difference 10
FT /note= "D-form residue"
FT Modified-site 12
FT /note= "C-terminal amide"
XX
PN WO9731019-A2.
XX
PD 28-AUG-1997.
XX
PF 20-FEB-1997; 97WO-IL00066.

XX
PR 22-FEB-1996; 96IL-0117223.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer
XX
PS Claim 21; Page 39; 80pp; English.
XX
CC This sequence represents a Leu/Lys diastereomer peptide of the
CC invention. The peptides of the invention have: (a) cytolytic activity on
CC pathogenic cells (pathogens and malignant cells not naturally present in
CC the body); but (b) no haemolytic activity, or such activity only at a
CC concentration significantly higher than that at which they lyse
CC pathogens. The peptides, their complexes and mixtures are used to treat
CC infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC or cancer, in human and veterinary medicine. Also, they can be used as
CC preservatives for food, cosmetics and agricultural produce, or as
CC agricultural pesticides. The absence of haemolytic activity (associated
CC with disturbance of alpha-helical structures) means that the peptides
CC have few if any toxic effects, and those that include D-aa will have
CC increased resistance to proteolytic degradation. Non-haemolytic,
CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC activity, allowing selection of agents for particular applications. Since
CC these random copolymers induce total lysis of bacterial cell walls,
CC resistance to them is unlikely to develop.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 18; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllklllkllk 12

RESULT 2
AAW35152
ID AAW35152 standard; peptide; 12 AA.
XX
AC AAW35152;
XX
DT 14-APR-1998 (first entry)
XX
DE Leu/Lys diastereomer peptide [D]-K1,5,9,12L2,6,7,11-K4L8.
XX
KW Leu/Lys diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW Honey bee venom; pardaxin; cytolytic activity; cancer;
KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW agricultural pesticide; cell wall lysis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "D-form residue"
FT Misc-difference 2
FT /note= "D-form residue"
FT Misc-difference 5
FT /note= "D-form residue"
FT Misc-difference 6
FT /note= "D-form residue"
FT Misc-difference 7
FT /note= "D-form residue"
FT Misc-difference 9
FT /note= "D-form residue"
FT Misc-difference 11
FT /note= "D-form residue"
FT Misc-difference 12
FT /note= "D-form residue"
FT Modified-site 12
FT /note= "C-terminal amide"
XX
PN WO9731019-A2.
XX
PD 28-AUG-1997.
XX
PF 20-FEB-1997; 97WO-IL00066.
XX
PR 22-FEB-1996; 96IL-0117223.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer
XX
PS Claim 21; Page 40; 80pp; English.
XX
CC This sequence represents a Leu/Lys diastereomer peptide of the
CC invention. The peptides of the invention have: (a) cytolytic activity on
CC pathogenic cells (pathogens and malignant cells not naturally present in
CC the body); but (b) no haemolytic activity, or such activity only at a
CC concentration significantly higher than that at which they lyse
CC pathogens. The peptides, their complexes and mixtures are used to treat
CC infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC or cancer, in human and veterinary medicine. Also, they can be used as
CC preservatives for food, cosmetics and agricultural produce, or as
CC agricultural pesticides. The absence of haemolytic activity (associated
CC with disturbance of alpha-helical structures) means that the peptides
CC have few if any toxic effects, and those that include D-aa will have
CC increased resistance to proteolytic degradation. Non-haemolytic,
CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC activity, allowing selection of agents for particular applications. Since
CC these random copolymers induce total lysis of bacterial cell walls,
CC resistance to them is unlikely to develop.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 18; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllklllkllk 12

RESULT 3
AAW82847
ID AAW82847 standard; peptide; 12 AA.
XX
AC AAW82847;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX

OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 12; Page 105; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllklllkllk 12

RESULT 4
AAW82850
ID AAW82850 standard; peptide; 12 AA.
XX
AC AAW82850;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 13; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllklllkllk 12

RESULT 5
AAW82856
ID AAW82856 standard; peptide; 12 AA.
XX
AC AAW82856;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX

PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 14; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllkllklllk 12

RESULT 6
AAB17413
ID AAB17413 standard; Peptide; 12 AA.
XX
AC AAB17413;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:517.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 378; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllkllklllk 12

RESULT 7
AAB17416
ID AAB17416 standard; Peptide; 12 AA.
XX
AC AAB17416;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:520.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.

XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 379; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllkllkllk 12

RESULT 8
AAB17483
ID AAB17483 standard; Peptide; 12 AA.
XX
AC AAB17483;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:587.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 401; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllkllkllk 12

RESULT 9
AAB17485
ID AAB17485 standard; Peptide; 12 AA.
XX
AC AAB17485;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:589.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;

XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 402; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 52; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllklllkllk 12

RESULT 10
AAW35231
ID AAW35231 standard; peptide; 13 AA.
XX
AC AAW35231;
XX
DT 14-APR-1998 (first entry)
XX
DE Diastereomer peptide [D]-L3,4,8,10-K4L8C.
XX
KW Diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW Honey bee venom; pardaxin; cytolytic activity; cancer;
KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW agricultural pesticide; cell wall lysis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 3
FT /note= "D-form residue"
FT Misc-difference 4
FT /note= "D-form residue"
FT Misc-difference 8
FT /note= "D-form residue"
FT Misc-difference 10
FT /note= "D-form residue"
XX
PN WO9731019-A2.
XX
PD 28-AUG-1997.
XX
PF 20-FEB-1997; 97WO-IL00066.
XX
PR 22-FEB-1996; 96IL-0117223.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer
XX
PS Example 7; Page 49; 80pp; English.
XX
CC This sequence represents a diastereomer peptide of the invention. This
CC sequence is used in a "bundle sequence", which is created by binding 5
CC copies of this sequence to peptide 23 (see AAW35149). The peptides of
CC the invention have: (a) cytolytic activity on pathogenic cells (pathogens
CC and malignant cells not naturally present in the body); but (b) no
CC haemolytic activity, or such activity only at a concentration
CC significantly higher than that at which they lyse pathogens. The
CC peptides, their complexes and mixtures are used to treat infections
CC (caused by bacteria, fungi, protozoa, mycoplasma or viruses) or cancer,
CC in human and veterinary medicine. Also, they can be used as preservatives
CC for food, cosmetics and agricultural produce, or as agricultural
CC pesticides. The absence of haemolytic activity (associated with
CC disturbance of alpha-helical structures) means that the peptides have few
CC if any toxic effects, and those that include D-aa will have increased
CC resistance to proteolytic degradation. Non-haemolytic, cytotoxic random
CC copolymers of pardaxin, each has a specific spectrum of activity,
CC allowing selection of agents for particular applications. Since these
CC random copolymers induce total lysis of bacterial cell walls, resistance
CC to them is unlikely to develop.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 52; DB 18; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllklllkllk 12

RESULT 11
AAB17482
ID AAB17482 standard; Peptide; 13 AA.
XX
AC AAB17482;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:586.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.

PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 401; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 52; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllklllkllk 12

RESULT 12
AAW82854
ID AAW82854 standard; peptide; 14 AA.
XX
AC AAW82854;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 14; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 52; DB 19; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 2 klllklllkllk 13

RESULT 13
AAW77378
ID AAW77378 standard; peptide; 37 AA.
XX
AC AAW77378;
XX
DT 14-DEC-1998 (first entry)
XX
DE Lytic peptide with alterable function 3.
XX
KW Biologically active peptide; hormone; drug; toxin;
KW lipid bilayer membrane; microorganism; parasite; virus.
XX
OS Synthetic.
XX
PN WO9841535-A2.
XX
PD 24-SEP-1998.
XX
PF 18-MAR-1998; 98WO-GB00799.
XX
PR 18-MAR-1997; 97GB-0005519.
XX
PA (ANMA-) ANMAT TECHNOLOGY LTD.
XX
PI Ajoula HS, Clarke DJ;
XX
DR WPI; 1998-521161/44.
XX
PT New modified peptide(s) - obtained by substitution with an amino
PT acid which is modifiable by a reaction and replacing other amino
PT acids which are not to be modified
XX

PS Claim 7; Page 22; 33pp; English.
XX
CC The peptides AAW77376-W77390 can be modified by the method of the
CC invention by substituting at least one amino acid of the peptide to
CC provide a peptide having at least one amino acid which is modifiable by
CC a reaction and replacing other amino acids in the peptide with amino
CC acids which are not modifiable by the reaction. The methods can be used
CC for the modification of biologically active peptides such as hormones,
CC drugs, toxins and peptides which act on lipid bilayer membranes. The
CC modified peptides can be used e.g. in the body of an animal or plant or
CC parts in order to affect the structure or integrity or permeability of a
CC foreign body such as a microorganism, parasite or virus present in the
CC body of the animal or plant or within the cells of the body of the animal
CC or plant.
XX
SQ Sequence 37 AA;

Query Match 100.0%; Score 52; DB 19; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 11 klllkllkllk 22

RESULT 14
AAW82858
ID AAW82858 standard; peptide; 77 AA.
XX
AC AAW82858;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 17; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 77 AA;

Query Match 100.0%; Score 52; DB 19; Length 77;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
||||||||||||
Db 1 klllkllkllk 12

RESULT 15
AAW82859
ID AAW82859 standard; peptide; 77 AA.
XX
AC AAW82859;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 17; Page 107; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be

CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 77 AA;

Query Match 100.0%; Score 52; DB 19; Length 77;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 KLLLKLLLKLLK 12
           ||||||||||||
Db      66 klllklllkllk 77
```

Search completed: June 17, 2002, 12:41:21
Job time: 296 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:38:45 ; Search time 46.42 Seconds
(without alignments)
24.840 Million cell updates/sec

Title: US-09-367-714A-23
Perfect score: 52
Sequence: 1 KLLLKLLLKLLK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	76.9	2513	2	G96536	hypothetical prote
2	37	71.2	137	2	A96914	uncharacterized pr
3	37	71.2	238	2	E71375	probable ABC trans
4	37	71.2	255	2	A60637	merozoite antigen
5	37	71.2	465	2	T30155	hypothetical prote
6	36	69.2	143	2	S03747	small membrane pro
7	36	69.2	1896	2	T01490	hypothetical prote
8	35	67.3	39	2	G85837	hypothetical prote
9	35	67.3	318	2	C81386	probable integral
10	35	67.3	319	2	A70102	conserved hypothet
11	35	67.3	3268	2	S69625	hypothetical prote
12	34	65.4	53	2	T03171	probable antibioti
13	34	65.4	191	2	F90392	hypothetical prote
14	34	65.4	213	2	T01464	hypothetical prote
15	34	65.4	235	2	I40627	probable transcrip
16	34	65.4	235	2	B97109	DNA-dependent RNA
17	34	65.4	282	2	G71932	hypothetical prote
18	34	65.4	433	2	A69735	phage PBSX termina
19	34	65.4	552	2	F71132	probable nitrite r
20	34	65.4	662	2	E95105	ABC transporter, p
21	34	65.4	662	2	E97973	hypothetical prote
22	34	65.4	707	2	H82709	colicin V secretio
23	34	65.4	984	1	DJNVCP	DNA-directed DNA p
24	34	65.4	986	2	T41809	DNA polymerase orf
25	34	65.4	1712	2	C71618	hypothetical prote
26	33	63.5	84	2	E96916	hypothetical prote
27	33	63.5	91	2	G70155	phosphocarrier pro
28	33	63.5	109	2	S42121	RNaseP C5 chain -
29	33	63.5	144	2	T18867	hypothetical prote
30	33	63.5	187	2	E95056	conserved hypothet
31	33	63.5	191	2	H72767	hypothetical prote
32	33	63.5	192	2	A97926	conserved hypothet
33	33	63.5	223	2	A99926	hypothetical prote
34	33	63.5	287	2	F72307	conserved hypothet
35	33	63.5	296	2	G97799	hypothetical prote
36	33	63.5	319	2	D90589	hypothetical prote
37	33	63.5	330	2	AH2188	hypothetical prote
38	33	63.5	333	1	DCRTDM	adenosylmethionine
39	33	63.5	334	1	DCHUDM	adenosylmethionine
40	33	63.5	334	1	DCHYDM	adenosylmethionine
41	33	63.5	334	2	A55948	adenosylmethionine
42	33	63.5	364	2	B36313	hypothetical 42K p
43	33	63.5	384	2	S66758	probable membrane
44	33	63.5	425	2	A70394	hypothetical prote
45	33	63.5	476	2	C84687	probable fatty aci

ALIGNMENTS

RESULT 1
G96536
hypothetical protein F2J10.9 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: G96536
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: G96536
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2513 <STO>
A;Cross-references: GB:AE005173; NID:g8569097; PIDN:AAF76442.1; GSPDB:GN00141
C;Genetics:
A;Gene: F2J10.9
A;Map position: 1

Query Match 76.9%; Score 40; DB 2; Length 2513;
Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2 LLLKLLLKLLK 12
           ||| |||||||
Db     797 LLLGLLLKLLK 807
```

RESULT 2
A96914
uncharacterized protein, Yje/RRF2 family [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: A96914
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L .; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: A96914
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-137 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK78100.1; PID:g15022941; GSPDB:GN00168

A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0115

Query Match 71.2%; Score 37; DB 2; Length 137;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLLK 12
: ||||| ||:|
Db 40 RFLLKLLRKLIK 51

RESULT 3
E71375
probable ABC transporter, ATP-binding protein - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 17-Mar-2000
C;Accession: E71375
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770
A;Accession: E71375
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-238 <COL>
A;Cross-references: GB:AE001188; GB:AE000520; NID:g3322282; PIDN:AAC65030.1; PID:g332229
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0035
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C;Keywords: ATP
F;27-207/Domain: ATP-binding cassette homology <ABC>

Query Match 71.2%; Score 37; DB 2; Length 238;
Best Local Similarity 80.0%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LLKLLLKLLK 12
|:||:|||||
Db 53 LVKLVLKLLK 62

RESULT 4
A60637
merozoite antigen LPMC-61 - Eimeria tenella (fragment)
C;Species: Eimeria tenella
C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 07-May-1999
C;Accession: A60637
R;Ko, C.; Smith II, C.K.; McDonell, M.
Mol. Biochem. Parasitol. 41, 53-64, 1990
A;Title: Identification and characterization of a target antigen of a monoclonal antibod
A;Reference number: A60637; MUID:90348718
A;Accession: A60637
A;Molecule type: mRNA
A;Residues: 1-255 <KOA>
A;Cross-references: GB:M30933
C;Keywords: tandem repeat
F;18-240/Region: glutamine-rich repeats

Query Match 71.2%; Score 37; DB 2; Length 255;
Best Local Similarity 81.8%; Pred. No. 44;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLLKLLLKLL 11
:|||||| ||
Db 2 RLLLKLLLLLL 12

RESULT 5
T30155
hypothetical protein C37A2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T30155
R;Le, T.T.; Kemp, K.; Scheet, P.
submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid C37A2.
A;Reference number: Z20746
A;Accession: T30155
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-465 <LET>
A;Cross-references: EMBL:U97194; PIDN:AAB52449.1; GSPDB:GN00019; CESP:C37A2.5
A;Experimental source: strain Bristol N2; clone C37A2
C;Genetics:
A;Gene: CESP:C37A2.5
A;Map position: 1
A;Introns: 47/1; 117/1; 185/1; 264/3; 364/2; 426/2

Query Match 71.2%; Score 37; DB 2; Length 465;
Best Local Similarity 70.0%; Pred. No. 77;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LLKLLLKLLK 12
|||::|||:||
Db 452 LLKIILKILK 461

RESULT 6
S03747
small membrane protein eag - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 21-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 20-Jun-2000
C;Accession: S03747; C69619
R;Perego, M.; Hoch, J.A.
Mol. Microbiol. 1, 125-132, 1987
A;Title: Isolation and sequence of the spo0E gene: its role in initiation of sporulat
A;Reference number: S03746; MUID:88260878
A;Accession: S03747
A;Molecule type: DNA
A;Residues: 1-143 <PER>
A;Cross-references: EMBL:Y00526; NID:g40181; PIDN:CAA68584.1; PID:g40183
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033
A;Accession: C69619
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-143 <KUN>
A;Cross-references: GB:Z99111; GB:AL009126; NID:g2633699; PIDN:CAB13238.1; PID:g26337
A;Experimental source: strain 168
C;Genetics:
A;Gene: eag
C;Superfamily: Bacillus subtilis small membrane protein eag

```
  Query Match             69.2%;  Score 36;  DB 2;  Length 143;
  Best Local Similarity   63.6%;  Pred. No. 38;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LLLKLLLKLLK 12
          ||||:||: :|
Db    114 LLLKMLLRFIK 124


RESULT   7
T01490
hypothetical protein F17O7.14 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Oct-1999
C;Accession: T01490
R;Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; Li,
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N
submitted to the EMBL Data Library, June 1998
A;Description: Arabidopsis thaliana chromosome 1 BAC F17O7 sequence.
A;Reference number: Z14334
A;Accession: T01490
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1896 <VYS>
A;Cross-references: EMBL:AC003671; NID:g2833627; PID:g3176689; GSPDB:GN00059; ATSP:F17O7
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: ATSP:F17O7.14
A;Map position: 1
A;Introns: 11/3; 43/3; 112/3; 1803/3


  Query Match             69.2%;  Score 36;  DB 2;  Length 1896;
  Best Local Similarity   72.7%;  Pred. No. 4.1e+02;
  Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LLLKLLLKLLK 12
          |||:|:|| ||
Db   1832 LLLELMLKTLK 1842


RESULT   8
G85837
hypothetical protein Z3270 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: G85837
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85837
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-39 <STO>
A;Cross-references: GB:AE005174; NID:g12516312; PIDN:AAG57163.1; GSPDB:GN00145; UWGP:Z32
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3270


  Query Match             67.3%;  Score 35;  DB 2;  Length 39;
  Best Local Similarity   54.5%;  Pred. No. 17;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LLLKLLLKLLK 12
          :|||||::|: |
Db      1 MLLKLIIKIFK 11


RESULT   9
C81386
probable integral membrane protein Cj0421c [imported] - Campylobacter jejuni (strain
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C;Accession: C81386
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chil
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A;Reference number: A81250; MUID:20150912
A;Accession: C81386
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-318 <PAR>
A;Cross-references: GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CAB74257.1; PID:g696
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj0421c


  Query Match             67.3%;  Score 35;  DB 2;  Length 318;
  Best Local Similarity   58.3%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLLKLLK 12
          ||:|| : |:||
Db    233 KLILKQIAKILK 244


RESULT  10
A70102
conserved hypothetical integral membrane protein BB0017 - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999
C;Accession: A70102
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943
A;Accession: A70102
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-319 <KLE>
A;Cross-references: GB:AE001116; GB:AE000783; NID:g2687896; PIDN:AAC66414.1; PID:g268
A;Experimental source: strain B31
C;Superfamily: conserved hypothetical protein yitT


  Query Match             67.3%;  Score 35;  DB 2;  Length 319;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLLKLLK 12
          |||| :|:| ||
Db     17 KLLLLILIKKLK 28


RESULT  11
S69625
hypothetical protein YDR457w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C;Accession: S69625
R;Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A;Description: The sequence of S. cerevisiae cosmids 9410, 8035, 8166, and 9787.
A;Reference number: S69554
A;Accession: S69625
A;Molecule type: DNA
A;Residues: 1-3268 <DIE>
```

```
A,Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64910.1; PID:g927738; MIPS:YDR457w
C,Genetics:
A,Gene: SGD:TOM1
A,Cross-references: SGD:S0002865; MIPS:YDR457w
A,Map position: 4R


  Query Match             67.3%;  Score 35;  DB 2;  Length 3268;
  Best Local Similarity   88.9%;  Pred. No. 9.9e+02;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLKLLLKLL 11
           ||||:|||| |
Db     515 LLKLVLKLL 523


RESULT  12
T03171
probable antibiotic polypeptide - Chilo iridescent virus
C;Species: Chilo iridescent virus
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 20-Aug-1999
C;Accession: T03171
R;Bahr, U.; Tidona, C.A.; Darai, G.
Virus Genes 15, 235-245, 1997
A;Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.101
A;Reference number: Z14834; MUID:98141693
A;Accession: T03171
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-53 <BAH>
A;Cross-references: EMBL:AF003534; NID:g2738385; PIDN:AAB94469.1; PID:g2738442
C;Superfamily: sillucin


  Query Match             65.4%;  Score 34;  DB 2;  Length 53;
  Best Local Similarity   70.0%;  Pred. No. 33;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLL 11
           :|||:| |||
Db       3 ILLKILFKLL 12


RESULT  13
F90392
hypothetical protein SSO2227 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: F90392
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: F90392
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-191 <KUR>
A;Cross-references: GB:AE006641; NID:g13815527; PIDN:AAK42397.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO2227


  Query Match             65.4%;  Score 34;  DB 2;  Length 191;
  Best Local Similarity   88.9%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 LKLLLKLLK 12
           || ||||||
Db      35 LKALLKLLK 43


RESULT  14
T01464
hypothetical protein T24H24.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 24-Mar-1999
C;Accession: T01464
R;Courtney, L.; Stoneking, T.; Langston, Y.; Mead, K.
submitted to the EMBL Data Library, August 1998
A;Description: The sequence of A. thaliana T24H24.
A;Reference number: Z14333
A;Accession: T01464
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-213 <COU>
A;Cross-references: EMBL:AF075598; NID:g3293581; PID:g3377838
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 48/1; 102/3
A;Note: T24H24.20


  Query Match             65.4%;  Score 34;  DB 2;  Length 213;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           :|:||||| | |
Db     200 QLILKLLLPLSK 211


RESULT  15
I40627
probable transcription initiation factor sigma E - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999
C;Accession: I40627; S34309
R;Wong, J.; Sass, C.; Bennett, G.N.
Gene 153, 89-92, 1995
A;Title: Sequence and arrangement of genes encoding sigma factors in Clostridium acet
A;Reference number: I40626; MUID:95189110
A;Accession: I40627
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-235 <RES>
A;Cross-references: EMBL:U07420; NID:g705344; PIDN:AAC43309.1; PID:g460971
R;Sauer, U.; Treuner, A.; Buchholz, M.; Duerre, P.
submitted to the EMBL Data Library, June 1993
A;Description: Sigma factor homologous genes in C. acetobutylicum.
A;Reference number: S34306
A;Accession: S34309
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 46-146,'P',148-235 <SAU>
A;Cross-references: EMBL:Z23079
C;Genetics:
A;Gene: sigE
C;Superfamily: transcription initiation factor sigma K; transcription initiation fact
C;Keywords: DNA binding; sigma factor; transcription initiation
F;60-235/Domain: transcription initiation factor sigma katF homology <KTF>


  Query Match             65.4%;  Score 34;  DB 2;  Length 235;
  Best Local Similarity   72.7%;  Pred. No. 1.3e+02;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLL 11
           | ||:|| |||
Db       2 KFLLRLLSKLL 12


Search completed: June 17, 2002, 12:42:57
```

Job time: 252 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:39:45 ; Search time 21.35 Seconds
(without alignments)
21.763 Million cell updates/sec

Title: US-09-367-714A-23
Perfect score: 52
Sequence: 1 KLLLKLLLKLLK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	71.2	238	1	Y035_TREPA	O83078 treponema p
2	37	71.2	255	1	LP61_EIMTE	P15714 eimeria ten
3	36	69.2	143	1	EAG_BACSU	P06630 bacillus su
4	35	67.3	1997	1	OTOF_HUMAN	Q9hc10 homo sapien
5	34	65.4	235	1	RPSE_CLOAB	P33657 clostridium
6	34	65.4	433	1	XTMB_BACSU	P39786 bacillus su
7	34	65.4	984	1	DPOL_NPVAC	P18131 autographa
8	34	65.4	986	1	DPOL_NPVBM	P41712 bombyx mori
9	33	63.5	31	1	LPL_BUCRP	Q53017 buchnera ap
10	33	63.5	109	1	RNPA_MYCCA	P43039 mycoplasma
11	33	63.5	333	1	DCAM_RAT	P17708 rattus norv
12	33	63.5	334	1	DCAM_BOVIN	P50243 bos taurus
13	33	63.5	334	1	DCAM_HUMAN	P17707 homo sapien
14	33	63.5	334	1	DCAM_MESAU	P28918 mesocricetu
15	33	63.5	334	1	DCM1_MOUSE	P31154 mus musculu
16	33	63.5	334	1	DCM2_MOUSE	P82184 mus musculu
17	33	63.5	334	1	DCM2_MUSSP	P82185 mus spretus
18	33	63.5	940	1	GBR2_RAT	O88871 rattus norv
19	32	61.5	201	1	EFA4_HUMAN	P52798 homo sapien
20	32	61.5	209	1	IL6_PHOVI	Q28819 phoca vitul
21	32	61.5	318	1	BST1_HUMAN	Q10588 homo sapien
22	32	61.5	386	1	RMAR_HANWI	P48849 hansenula w
23	32	61.5	402	1	SHBG_HUMAN	P04278 homo sapien
24	32	61.5	475	1	ASPA_BACSU	P26899 bacillus su
25	32	61.5	483	1	VATH_BOVIN	O46563 bos taurus
26	32	61.5	483	1	VATH_HUMAN	Q9ui12 homo sapien
27	32	61.5	483	1	VATH_PIG	Q9tvc1 sus scrofa
28	32	61.5	557	1	FACC_RAT	O35870 rattus norv
29	32	61.5	630	1	Y242_MYCGE	P47484 mycoplasma
30	32	61.5	743	1	PURL_BACHD	Q9kf57 bacillus ha
31	32	61.5	811	1	RFXL_YEAST	P48743 saccharomyc
32	32	61.5	875	1	ENV_BIV06	P19556 bovine immu
33	32	61.5	904	1	ENV_BIV27	P19557 bovine immu
34	32	61.5	1472	1	A2MG_RAT	P06238 rattus norv
35	32	61.5	1941	1	UBR1_KLULA	O60014 kluyveromyc
36	32	61.5	2124	1	Y192_HUMAN	Q93074 homo sapien
37	32	61.5	2470	1	TOR1_YEAST	P35169 saccharomyc
38	32	61.5	3744	1	YHP9_YEAST	P38811 saccharomyc
39	31	59.6	159	1	YIB6_YEAST	P40548 saccharomyc
40	31	59.6	193	1	INF1_CHICK	P42165 gallus gall
41	31	59.6	193	1	INF3_CHICK	Q90872 gallus gall
42	31	59.6	229	1	SOML_TETMU	Q9i9h4 tetraodon m
43	31	59.6	231	1	SOM1_SPAAU	P54863 sparus aura
44	31	59.6	231	1	SOM2_SPAAU	P79894 sparus aura
45	31	59.6	231	1	SOML_SCIOC	Q9ygk7 sciaenops o

ALIGNMENTS

RESULT 1
Y035_TREPA
ID Y035_TREPA STANDARD; PRT; 238 AA.
AC O83078;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable metal transport system ATP-binding protein TP0035.
GN TP0035.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -!- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM
CC TP0034/TP0035/TP0036 FOR A METAL. PROBABLY RESPONSIBLE FOR ENERGY
CC COUPLING TO THE TRANSPORT SYSTEM.
CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC ---

CC ---
DR EMBL; AE001188; AAC65030.1; -.
DR TIGR; TP0035; -.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; FALSE_NEG.
KW Hypothetical protein; Transport; Inner membrane; ATP-binding;
KW Complete proteome.
FT NP_BIND 44 51 ATP (POTENTIAL).
SQ SEQUENCE 238 AA; 26460 MW; 673E7B4882BE4D29 CRC64;

Query Match 71.2%; Score 37; DB 1; Length 238;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3 LLKLLLKLLK 12
           |:||:|||||
Db      53 LVKLVLKLLK 62


RESULT   2
LP61_EIMTE
ID   LP61_EIMTE      STANDARD;      PRT;   255 AA.
AC   P15714;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Antigen LPMC-61 (Fragment).
OS   Eimeria tenella.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
OC   Eimeria.
OX   NCBI_TaxID=5802;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Sporozoite;
RX   MEDLINE=90348718; PubMed=2200963;
RA   Ko C., Smith C.K. II, McDonell M.;
RT   "Identification and characterization of a target antigen of a
RT   monoclonal antibody directed against Eimeria tenella merozoites.";
RL   Mol. Biochem. Parasitol. 41:53-64(1990).
CC   -!- FUNCTION: UNKNOWN. THE GLN-RICH TANDEM REPEATS MAY BE IMPORTANT
CC       FOR AN UNKNOWN ASPECT OF THE PARASITIC LIFE CYCLE. MAY BE AN
CC       IMPORTANT IMMUNOGEN.
CC   -!- SUBUNIT: MAY BE COVALENTLY LINKED BY DISULFIDE BONDS TO OTHER
CC       POLYPEPTIDES TO FORM THE 80 kDa ANTIGEN.
CC   -!- DEVELOPMENTAL STAGE: PRESENT IN ALL STAGES THROUGHOUT THE
CC       SPORULATION OF THE OOCYSTS AND IN THE SPOROZOITES FOLLOWING
CC       EXCYSTATION.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30933; AAA29079.1; -.
DR   PIR; A60637; A60637.
KW   Antigen; Sporozoite; Repeat; Sporulation.
FT   NON_TER       1      1
FT   DOMAIN       18    210       12 X APPROXIMATE TANDEM REPEATS, GLN-
FT                                 RICH.
FT   REPEAT       18     48       1.
FT   REPEAT       49     57       2.
FT   REPEAT       58     65       3.
FT   REPEAT       66     78       4.
FT   REPEAT       79     90       5.
FT   REPEAT       91    103       6.
FT   REPEAT      104    140       7.
FT   REPEAT      141    152       8.
FT   REPEAT      153    164       9.
FT   REPEAT      165    172       10.
FT   REPEAT      173    192       11.
FT   REPEAT      193    210       12.
FT   NON_TER     255    255
SQ   SEQUENCE   255 AA;  31267 MW;  8C5E6005FFFC2DB3 CRC64;


  Query Match             71.2%;  Score 37;  DB 1;  Length 255;
  Best Local Similarity   81.8%;  Pred. No. 16;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLL 11
           :|||||||  ||
Db       2 RLLLKLLLLLL 12


RESULT   3
EAG_BACSU
ID   EAG_BACSU      STANDARD;      PRT;   143 AA.
AC   P06630;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 16.4 kDa protein in SPO0E 3'region.
GN   EAG.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=88260878; PubMed=2838724;
RA   Perego M., Hoch J.A.;
RT   "Isolation and sequence of the spo0E gene: its role in initiation of
RT   sporulation in Bacillus subtilis.";
RL   Mol. Microbiol. 1:125-132(1987).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y00526; CAA68584.1; -.
DR   EMBL; Z99111; CAB13238.1; -.
DR   PIR; S03747; S03747.
DR   SubtiList; BG10770; eag.
KW   Hypothetical protein; Sporulation; Complete proteome.
SQ   SEQUENCE   143 AA;  16429 MW;  D7410B50963D7A75 CRC64;


  Query Match             69.2%;  Score 36;  DB 1;  Length 143;
  Best Local Similarity   63.6%;  Pred. No. 14;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLLK 12
           ||||:||: :|
Db     114 LLLKMLLRFIK 124


RESULT   4
OTOF_HUMAN
ID   OTOF_HUMAN      STANDARD;      PRT;  1997 AA.
AC   Q9HC10; Q9HC09; Q9Y650; Q9HC08;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Otoferlin (Fer-1 like protein 2).
GN   OTOF OR FER1L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain;
RX   MEDLINE=20395831; PubMed=10903124;
RA   Yasunaga S., Grati M., Chardenoux S., Smith T.N., Friedman T.B.,
RA   Lalwani A.K., Wilcox E.R., Petit C.;
RT   "OTOF encodes multiple long and short isoforms: genetic evidence that
RT   the long ones underlie recessive deafness DFNB9.";
RL   Am. J. Hum. Genet. 67:591-600(2000).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 4).
RC   TISSUE=Fetal;
RX   MEDLINE=99206603; PubMed=10192385;
```

```
RA   Yasunaga S., Grati M., Cohen-Salmon M., El-Amraoui A., Mustapha M.,
RA   Salem N., El-Zir E., Loiselet J., Petit C.;
RT   "A mutation in OTOF, encoding otoferlin, a FER-1-like protein, causes
RT   DFNB9, a nonsyndromic form of deafness.";
RL   Nat. Genet. 21:363-369(1999).
CC   -!- FUNCTION: Might be involved in the Ca(2+)-triggered synaptic
CC       vesicle-plasma membrane fusion.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC   -!- ALTERNATIVE PRODUCTS: at least 4 isoforms; 1/long (shown here),
CC       2/short-1, 3/short-2 and 4/short-3; are produced by alternative
CC       splicing.
CC   -!- TISSUE SPECIFICITY: Isoform 1 and isoform 3 are found in adult
CC       brain. Isoform 2 is expressed in the fetus and in adult brain,
CC       heart, placenta, skeletal muscle and kidney.
CC   -!- DISEASE: Defects in OTOF are the cause of nonsyndromic autosomal
CC       recessive deafness 9 (DFNB9).
CC   -!- SIMILARITY: BELONGS TO THE FERLIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 4 C2 DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF183185; AAG12991.1; -.
DR   EMBL; AF183186; AAG12992.1; -.
DR   EMBL; AF183187; AAG17468.1; -.
DR   EMBL; AF107403; AAD26117.1; -.
DR   HSSP; P04410; 1A25.
DR   MIM; 603681; -.
DR   MIM; 601071; -.
DR   InterPro; IPR000008; C2.
DR   Pfam; PF00168; C2; 4.
DR   PRINTS; PR00360; C2DOMAIN.
DR   SMART; SM00239; C2; 6.
DR   PROSITE; PS00499; C2_DOMAIN_1; 2.
DR   PROSITE; PS50004; C2_DOMAIN_2; 4.
KW   Transmembrane; Repeat; Alternative splicing; Deafness.
FT   DOMAIN        1   1963       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1964   1984       POTENTIAL.
FT   DOMAIN     1985   1997       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      241    338       C2 DOMAIN 1.
FT   DOMAIN      404    514       C2 DOMAIN 2.
FT   DOMAIN      947   1052       C2 DOMAIN 3.
FT   DOMAIN     1479   1577       C2 DOMAIN 4.
FT   DOMAIN     1303   1310       POLY-LYS.
FT   DOMAIN     1314   1320       POLY-GLU.
FT   DOMAIN     1965   1983       POLY-LEU.
FT   VARSPLIC      1    747       MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT   VARSPLIC   1245   1264       MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT   VARSPLIC      1    690       MISSING (IN ISOFORM 3).
FT   VARSPLIC    691    738       MRPQVTDRNYFHLPYLERKPCIYIKSWWPDQRRRLYNANIM
FT                                DHIADKL -> MMTDTQDGPSESSQIMRSLTLLINREEAFG
FT                                EAGEAGLWPSITHTPDSQ (IN ISOFORM 3).
FT   VARSPLIC   1943   1997       SFIWFLNPLKSARYFLWHTYRWLLLKLLLLLLLLLLLALFL
FT                                YSVPGYLVKKILGA -> AFVWFLNPLKSIKYLICTRYKWL
FT                                IIKIVLALLGLLMLGLFLYSLPGYMVKKLLGA (IN
FT                                ISOFORM 2).
FT   CONFLICT   1088   1088       P -> L (IN REF. 1; AAG12991).
SQ   SEQUENCE   1997 AA;  226735 MW;  39D10CB5220638AE CRC64;


  Query Match             67.3%;  Score 35;  DB 1;  Length 1997;
  Best Local Similarity   90.0%;  Pred. No. 2.4e+02;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLL 11
           ||||||| ||
Db    1965 LLLKLLLLLL 1974
```

```
RESULT   5
RPSE_CLOAB
ID   RPSE_CLOAB     STANDARD;      PRT;   235 AA.
AC   P33657;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   RNA polymerase sigma-E factor.
GN   SIGE OR CAC1695.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=95189110; PubMed=7883192;
RA   Wong J., Sass C., Bennett G.N.;
RT   "Sequence and arrangement of genes encoding sigma factors in
RT   Clostridium acetobutylicum ATCC 824.";
RL   Gene 153:89-92(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RA   Santangelo J.D., Kuhn A., Treuner A., Durre P.;
RT   "Sporulation and time course expression of sigma factor homologous
RT   genes in Clostridium acetobutylicum.";
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
RN   [4]
RP   SEQUENCE OF 46-235 FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=95050216; PubMed=7961408;
RA   Sauer U., Treuner A., Buchholz M., Santangelo J.D., Durre P.;
RT   "Sporulation and primary sigma factor homologous genes in Clostridium
RT   acetobutylicum.";
RL   J. Bacteriol. 176:6572-6582(1994).
CC   -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC       ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC       THEN IS RELEASED. THIS SIGMA FACTOR IS RESPONSIBLE FOR THE
CC       EXPRESSION OF SPORULATION SPECIFIC GENES (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U07420; AAC43309.1; -.
DR   EMBL; Z23079; CAA80617.1; -.
DR   EMBL; AE007679; AAK79661.1; -.
DR   PIR; S34309; S34309.
DR   HSSP; P00579; 1SIG.
DR   InterPro; IPR000943; Sigma_70.
DR   Pfam; PF00140; sigma70; 1.
DR   PROSITE; PS00715; SIGMA70_1; 1.
DR   PROSITE; PS00716; SIGMA70_2; 1.
KW   Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
KW   DNA-binding; Sporulation; Complete proteome.
```

```
FT   DOMAIN       82     95       POLYMERASE CORE BINDING (POTENTIAL).
FT   DNA_BIND     202    221      H-T-H MOTIF (BY SIMILARITY).
FT   CONFLICT     147    147      L -> P (IN REF. 4).
SQ   SEQUENCE   235 AA;  26969 MW;  C726E18E6C93A903 CRC64;


  Query Match             65.4%;  Score 34;  DB 1;  Length 235;
  Best Local Similarity   72.7%;  Pred. No. 48;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLL 11
             | ||:|| |||
Db       2 KFLLRLLSKLL 12


RESULT   6
XTMB_BACSU
ID   XTMB_BACSU     STANDARD;      PRT;   433 AA.
AC   P39786;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   PBSX phage terminase large subunit.
GN   XTMB.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Krogh S., O'Reilly M., Nolan N., Devine K.M.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-76 FROM N.A.
RC   STRAIN=168 / SO113;
RX   MEDLINE=94364963; PubMed=8083174;
RA   McDonnell G.E., Wood H., Devine K.M., McConnell D.J.;
RT   "Genetic control of bacterial suicide: regulation of the induction of
RT   PBSX in Bacillus subtilis.";
RL   J. Bacteriol. 176:5820-5830(1994).
CC   -!- FUNCTION: FUNCTION AS A TERMINASE.
CC   -!- SUBUNIT: DIMER OF A SMALL AND A LARGE SUBUNIT (POTENTIAL).
CC   -!- SIMILARITY: STRONG, TO B.SUBTILIS YQAT.
CC   -!- SIMILARITY: TO LARGE SUBUNIT OF B.SUBTILIS PHAGE SPP1 TERMINASE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z70177; CAA94059.1; -.
DR   EMBL; Z34287; CAA84048.1; -.
DR   EMBL; Z99110; CAB13115.1; -.
DR   PIR; S47115; S47115.
DR   SubtiList; BG11000; xtmB.
KW   DNA packaging; Complete proteome.
SQ   SEQUENCE   433 AA;  51150 MW;  471FC77DFEA2CA10 CRC64;


  Query Match             65.4%;  Score 34;  DB 1;  Length 433;
  Best Local Similarity   77.8%;  Pred. No. 85;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 LKLLLKLLK 12
             |::|||||
Db      38 LKIVLKLLK 46


RESULT   7
DPOL_NPVAC
ID   DPOL_NPVAC     STANDARD;      PRT;   984 AA.
AC   P18131;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   DNA polymerase (EC 2.7.7.7).
GN   POL.
OS   Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=46015;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L1;
RX   MEDLINE=89073763; PubMed=3059678;
RA   Tomalski M.D., Wu J.G., Miller L.K.;
RT   "The location, sequence, transcription, and regulation of a
RT   baculovirus DNA polymerase gene.";
RL   Virology 167:591-600(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C6;
RX   MEDLINE=94303173; PubMed=8030224;
RA   Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT   "The complete DNA sequence of Autographa californica nuclear
RT   polyhedrosis virus.";
RL   Virology 202:586-605(1994).
CC   -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC       + {DNA}(N).
CC   -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M20744; AAA46692.1; -.
DR   EMBL; L22858; AAA66695.1; -.
DR   PIR; A31832; DJNVCP.
DR   InterPro; IPR002064; DNA_pol_B.
DR   Pfam; PF00136; DNA_pol_B; 1.
DR   Pfam; PF03104; DNA_pol_B_exo; 1.
DR   PRINTS; PR00106; DNAPOLB.
DR   SMART; SM00486; POLBc; 1.
DR   PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW   Transferase; DNA-directed DNA polymerase; DNA replication;
KW   DNA-binding; Early protein.
FT   DOMAIN      724    727      POLY-LYS.
FT   DOMAIN      946    960      POLY-ASP.
FT   CONFLICT    830    830      R -> W (IN REF. 1).
SQ   SEQUENCE   984 AA;  114307 MW;  156ABB6BA1B45A21 CRC64;


  Query Match             65.4%;  Score 34;  DB 1;  Length 984;
  Best Local Similarity   70.0%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLL 11
             :::|||||||
Db     566 IVVKLLLKLL 575


RESULT   8
DPOL_NPVBM
ID   DPOL_NPVBM     STANDARD;      PRT;   986 AA.
AC   P41712; O92430;
DT   01-NOV-1995 (Rel. 32, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
```

```
DE   DNA polymerase (EC 2.7.7.7).
GN   POL.
OS   Bombyx mori nuclear polyhedrosis virus (BmNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=10458;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95133178; PubMed=7831799;
RA   Chaeychomsri S., Ikeda M., Kobayashi M.;
RT   "Nucleotide sequence and transcriptional analysis of the DNA
RT   polymerase gene of Bombyx mori nuclear polyhedrosis virus.";
RL   Virology 206:435-447(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=T3;
RA   Gomi S., Majima K., Maeda S.;
RT   "Sequence analysis of the genome of Bombyx mori
RT   nucleopolyhedrovirus.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC       + {DNA}(N).
CC   -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D16231; BAA03756.1; -.
DR   EMBL; L33180; AAC63738.1; -.
DR   InterPro; IPR002064; DNA_pol_B.
DR   Pfam; PF00136; DNA_pol_B; 1.
DR   Pfam; PF03104; DNA_pol_B_exo; 1.
DR   PRINTS; PR00106; DNAPOLB.
DR   SMART; SM00486; POLBc; 1.
DR   PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW   Transferase; DNA-directed DNA polymerase; DNA replication;
KW   DNA-binding; Early protein.
FT   DOMAIN      724    727       POLY-LYS.
FT   DOMAIN      947    951       POLY-ASP.
FT   DOMAIN      954    959       POLY-ASP.
FT   CONFLICT    116    116       A -> S (IN REF. 1).
FT   CONFLICT    245    245       H -> Y (IN REF. 1).
FT   CONFLICT    250    250       H -> Y (IN REF. 1).
FT   CONFLICT    258    258       V -> I (IN REF. 1).
FT   CONFLICT    478    479       TA -> AG (IN REF. 1).
FT   CONFLICT    941    941       S -> G (IN REF. 1).
FT   CONFLICT    952    952       N -> NDN (IN REF. 1).
SQ   SEQUENCE   986 AA;  114418 MW;  503E39F4A0BCC125 CRC64;


  Query Match             65.4%;  Score 34;  DB 1;  Length 986;
  Best Local Similarity   70.0%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LLLKLLLKLL 11
          :::|||||||
Db    566 IVVKLLLKLL 575


RESULT   9
LPL_BUCRP
ID   LPL_BUCRP      STANDARD;      PRT;    31 AA.
AC   Q53017;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Leu operon leader peptide.
GN   LEUL OR LEUO.
OS   Buchnera aphidicola (subsp. Rhopalosiphum padi).
OG   Plasmid pRPE.
OC   Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX   NCBI_TaxID=98793;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95333198; PubMed=7608990;
RA   Bracho A.M., Martinez-Torres D., Moya A., Latorre A.;
RT   "Discovery and molecular characterization of a plasmid localized in
RT   Buchnera sp. bacterial endosymbiont of the aphid Rhopalosiphum
RT   padi.";
RL   J. Mol. Evol. 41:67-73(1995).
CC   -!- FUNCTION: THIS PROTEIN IS INVOLVED IN CONTROL OF THE BIOSYNTHESIS
CC       OF LEUCINE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X71612; CAA50613.1; -.
KW   Leader peptide; Leucine biosynthesis; Plasmid.
SQ   SEQUENCE   31 AA;  3920 MW;  4D3E5E1C31C85413 CRC64;


  Query Match             63.5%;  Score 33;  DB 1;  Length 31;
  Best Local Similarity   81.8%;  Pred. No. 11;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLLKLL 11
            |||| ||| ||
Db      8 KLLLLLLLYLL 18


RESULT  10
RNPA_MYCCA
ID   RNPA_MYCCA     STANDARD;      PRT;   109 AA.
AC   P43039;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ribonuclease P protein component (EC 3.1.26.5) (RNaseP protein)
DE   (RNase P protein) (Protein C5).
GN   RNPA.
OS   Mycoplasma capricolum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Entomoplasmataceae.
OX   NCBI_TaxID=2095;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 27343;
RX   MEDLINE=94051609; PubMed=8233831;
RA   Miyata M., Sano K.-I., Okada R., Fukumura T.;
RT   "Mapping of replication initiation site in Mycoplasma capricolum
RT   genome by two-dimensional gel-electrophoretic analysis.";
RL   Nucleic Acids Res. 21:4816-4823(1993).
CC   -!- FUNCTION: RNaseP catalyzes the removal of the 5'-leader sequence
CC       from pre-tRNA to produce the mature 5'terminus. It can also
CC       cleave other RNA substrates such as 4.5S RNA. The protein
CC       component plays an auxiliary but essential role in vivo by binding
CC       to the 5'-leader sequence and broadening the substrate specificity
CC       of the ribozyme (By similarity).
CC   -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-
CC       extra-nucleotide from tRNA precursor.
CC   -!- SUBUNIT: Consists of a catalytic RNA component (M1 or rnpB) and a
CC       protein subunit (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RNPA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D14982; BAA03619.1; -.
DR   HSSP; P25814; 1A6F.
DR   InterPro; IPR000100; Ribonuclease_P.
DR   Pfam; PF00825; Ribonuclease_P; 1.
DR   PROSITE; PS00648; RIBONUCLEASE_P; 1.
KW   Hydrolase; Nuclease; Endonuclease; tRNA processing; RNA-binding.
SQ   SEQUENCE   109 AA;  12900 MW;  ACF520A0982C0D12 CRC64;


  Query Match             63.5%;  Score 33;  DB 1;  Length 109;
  Best Local Similarity   70.0%;  Pred. No. 34;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 LLKLLLKLLK 12
           | |||:||:|
Db      97 LSKLLIKLIK 106


RESULT  11
DCAM_RAT
ID   DCAM_RAT       STANDARD;      PRT;   333 AA.
AC   P17708;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50) (AdoMetDC)
DE   (SamDC) [Contains: S-adenosylmethionine decarboxylase alpha chain; S-
DE   adenosylmethionine decarboxylase beta chain].
GN   AMD1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90215298; PubMed=2323572;
RA   Pulkka A., Keraenen M.R., Salmela A., Salmikangas P., Ihalainen R.,
RA   Pajunen A.;
RT   "Nucleotide sequence of rat S-adenosylmethionine decarboxylase cDNA.
RT   Comparison with an intronless rat pseudogene.";
RL   Gene 86:193-199(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89034205; PubMed=2460457;
RA   Pajunen A., Crozat A., Jaenne O.A., Ihalainen R., Laitinen P.H.,
RA   Stanley B., Madhubala R., Pegg A.E.;
RT   "Structure and regulation of mammalian S-adenosylmethionine
RT   decarboxylase.";
RL   J. Biol. Chem. 263:17040-17049(1988).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92038054; PubMed=1936275;
RA   Pulkka A., Ihalainen R., Aatsinki J., Pajunen A.;
RT   "Structure and organization of the gene encoding rat S-
RT   adenosylmethionine decarboxylase.";
RL   FEBS Lett. 291:289-295(1991).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Spleen;
RX   MEDLINE=93300506; PubMed=8314573;
RA   Pulkka A., Ihalainen R., Suorsa A., Riviere M., Szpirer J.,
RA   Pajunen A.;
RT   "Structures and chromosomal localizations of two rat genes encoding
RT   S-adenosylmethionine decarboxylase.";
RL   Genomics 16:342-349(1993).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = (5-deoxy-5-
CC       adenosyl)(3-aminopropyl) methylsulfonium salt + CO(2).
CC   -!- COFACTOR: PYRUVOYL GROUP.
CC   -!- PATHWAY: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE
CC       AMINOPROPYL MOIETY REQUIRED FOR SPERMIDINE AND SPERMINE
CC       BIOSYNTHESIS FROM PUTRESCINE.
CC   -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC ADOMETDC FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M34464; AAA40683.1; -.
DR   EMBL; M64274; AAA42105.1; -.
DR   EMBL; Z15109; CAA78814.1; -.
DR   EMBL; Z15122; CAA78814.1; JOINED.
DR   EMBL; Z15123; CAA78814.1; JOINED.
DR   PIR; JQ0439; DCRTDM.
DR   PIR; S18487; S18487.
DR   HSSP; P17707; 1JEN.
DR   InterPro; IPR001985; SAM_decarbox.
DR   Pfam; PF01536; SAM_decarbox; 1.
DR   ProDom; PD002379; SAM_decarbox; 1.
DR   PROSITE; PS01336; ADOMETDC; 1.
KW   Spermidine biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen.
FT   CHAIN          1     67       S-ADENOSYLMETHIONINE DECARBOXYLASE BETA
FT                                 CHAIN.
FT   CHAIN         68    333       S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA
FT                                 CHAIN.
FT   SITE          67     68       CLEAVAGE (NONHYDROLYTIC).
FT   MOD_RES       68     68       CONVERTED TO A PYRUVOYL GROUP.
FT   ACT_SITE       8      8       IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                 SIMILARITY).
FT   ACT_SITE      11     11       IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                 SIMILARITY).
FT   ACT_SITE      82     82       IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                 SIMILARITY).
FT   CONFLICT       5      5       H -> P (IN REF. 2).
FT   CONFLICT     146    146       A -> G (IN REF. 1).
SQ   SEQUENCE   333 AA;  38137 MW;  9323E2D38BD8FEF1 CRC64;


  Query Match             63.5%;  Score 33;  DB 1;  Length 333;
  Best Local Similarity   72.7%;  Pred. No. 98;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLLK 12
           |||| |: |||
Db      86 LLLKALVPLLK 96


RESULT  12
DCAM_BOVIN
ID   DCAM_BOVIN     STANDARD;      PRT;   334 AA.
AC   P50243;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50) (AdoMetDC)
DE   (SamDC) [Contains: S-adenosylmethionine decarboxylase alpha chain; S-
DE   adenosylmethionine decarboxylase beta chain].
GN   AMD1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Hill J.R., Morris D.R.;
RL   Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 209-232 FROM N.A.
RX   MEDLINE=86304300; PubMed=3017942;
RA   Mach M., White M.W., Neubauer M., Degen J.L., Morris D.R.;
RT   "Isolation of a cDNA clone encoding S-adenosylmethionine
RT   decarboxylase. Expression of the gene in mitogen-activated
RT   lymphocytes.";
RL   J. Biol. Chem. 261:11697-11703(1986).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = (5-deoxy-5-
CC       adenosyl)(3-aminopropyl) methylsulfonium salt + CO(2).
CC   -!- COFACTOR: PYRUVOYL GROUP.
CC   -!- PATHWAY: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE
CC       AMINOPROPYL MOIETY REQUIRED FOR SPERMIDINE AND SPERMINE
CC       BIOSYNTHESIS FROM PUTRESCINE.
CC   -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC ADOMETDC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M95605; AAA30359.1; -.
DR   EMBL; M14289; AAA30360.1; -.
DR   HSSP; P17707; 1JEN.
DR   InterPro; IPR001985; SAM_decarbox.
DR   Pfam; PF01536; SAM_decarbox; 1.
DR   ProDom; PD002379; SAM_decarbox; 1.
DR   PROSITE; PS01336; ADOMETDC; 1.
KW   Spermidine biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen.
FT   CHAIN         1     67       S-ADENOSYLMETHIONINE DECARBOXYLASE BETA
FT                                CHAIN.
FT   CHAIN        68    334       S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA
FT                                CHAIN.
FT   SITE         67     68       CLEAVAGE (NONHYDROLYTIC).
FT   MOD_RES      68     68       CONVERTED TO A PYRUVOYL GROUP.
FT   ACT_SITE      8      8       IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                SIMILARITY).
FT   ACT_SITE     11     11       IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                SIMILARITY).
FT   ACT_SITE     82     82       IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                SIMILARITY).
SQ   SEQUENCE   334 AA;  38364 MW;  D8AFB98C9DD8A1E9 CRC64;


  Query Match           63.5%;  Score 33;  DB 1;  Length 334;
  Best Local Similarity  72.7%;  Pred. No. 98;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLLK 12
           |||| |: |||
Db      86 LLLKALVPLLK 96


RESULT 13
DCAM_HUMAN
ID   DCAM_HUMAN     STANDARD;      PRT;   334 AA.
AC   P17707; Q9BWK4;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50) (AdoMetDC)
DE   (SamDC) [Contains: S-adenosylmethionine decarboxylase alpha chain; S-
DE   adenosylmethionine decarboxylase beta chain].
GN   AMD1 OR AMD.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=89034205; PubMed=2460457;
RA   Pajunen A., Crozat A., Jaenne O.A., Ihalainen R., Laitinen P.H.,
RA   Stanley B., Madhubala R., Pegg A.E.;
RT   "Structure and regulation of mammalian S-adenosylmethionine
RT   decarboxylase.";
RL   J. Biol. Chem. 263:17040-17049(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Choriocarcinoma;
RA   Strausberg R.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   MUTAGENESIS.
RX   MEDLINE=92011599; PubMed=1917972;
RA   Stanley B.A., Pegg A.E.;
RT   "Amino acid residues necessary for putrescine stimulation of human S-
RT   adenosylmethionine decarboxylase proenzyme processing and catalytic
RT   activity.";
RL   J. Biol. Chem. 266:18502-18506(1991).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).
RX   MEDLINE=99306040; PubMed=10378277;
RA   Ekstrom J.L., Mathews I.I., Stanley B.A., Pegg A.E., Ealick S.E.;
RT   "The crystal structure of human S-adenosylmethionine decarboxylase at
RT   2.25-A resolution reveals a novel fold.";
RL   Structure 7:583-595(1999).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = (5-deoxy-5-
CC       adenosyl)(3-aminopropyl) methylsulfonium salt + CO(2).
CC   -!- COFACTOR: PYRUVOYL GROUP.
CC   -!- ENZYME REGULATION: BOTH PROENZYME PROCESSING AND CATALYTIC
CC       ACTIVITY ARE STIMULATED BY PUTRESCINE.
CC   -!- PATHWAY: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE
CC       AMINOPROPYL MOIETY REQUIRED FOR SPERMIDINE AND SPERMINE
CC       BIOSYNTHESIS FROM PUTRESCINE.
CC   -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS.
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC ADOMETDC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21154; AAA51716.1; -.
DR   EMBL; BC000171; AAH00171.1; -.
DR   PIR; A31786; DCHUDM.
DR   PDB; 1JEN; 01-JUN-99.
DR   MIM; 180980; -.
DR   InterPro; IPR001985; SAM_decarbox.
DR   Pfam; PF01536; SAM_decarbox; 1.
DR   ProDom; PD002379; SAM_decarbox; 1.
DR   PROSITE; PS01336; ADOMETDC; 1.
KW   Spermidine biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen;
KW   3D-structure.
FT   CHAIN         1     67       S-ADENOSYLMETHIONINE DECARBOXYLASE BETA
FT                                CHAIN.
FT   CHAIN        68    334       S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA
FT                                CHAIN.
FT   SITE         67     68       CLEAVAGE (NONHYDROLYTIC).
FT   MOD_RES      68     68       CONVERTED TO A PYRUVOYL GROUP.
FT   ACT_SITE      8      8       IMPORTANT FOR CATALYTIC ACTIVITY.
FT   ACT_SITE     11     11       IMPORTANT FOR CATALYTIC ACTIVITY, AND
FT                                PUTRESCINE STIMULAION OF PROCESSING.
FT   ACT_SITE     82     82       IMPORTANT FOR CATALYTIC ACTIVITY.
FT   MUTAGEN       8      8       E->Q: LOSS OF ACTIVITY. NORMAL
FT                                PUTRESCINE-STIMULATED PROCESSING.
FT   MUTAGEN      11     11       E->Q: LOSS OF ACTIVITY. LOSS OF
```

```
FT                                   PUTRESCINE-STIMULATED PROCESSING.
FT   MUTAGEN      15     15          E->Q: LITTLE EFFECT.
FT   MUTAGEN      49     49          C->A: LITTLE EFFECT.
FT   MUTAGEN      61     61          E->Q: LITTLE EFFECT.
FT   MUTAGEN      67     67          E->Q: LITTLE EFFECT.
FT   MUTAGEN      80     80          K->A: GREATLY REDUCED CATALYTIC
FT                                   ACTIVITY. NO PUTRESCINE-STIMULATED
FT                                   PROCESSING.
FT   MUTAGEN      82     82          C->A: LOSS OF ACTIVITY. GREATLY REDUCED
FT                                   PUTRESCINE-STIMULATED PROCESSING.
FT   MUTAGEN     226    226          C->A: LITTLE EFFECT.
FT   MUTAGEN     247    247          E->Q: LITTLE EFFECT.
FT   MUTAGEN     249    249          E->Q: LITTLE EFFECT.
FT   CONFLICT    146    146          G -> A (IN REF. 2).
SQ   SEQUENCE   334 AA;  38325 MW;  F78F93AAE28A92DC CRC64;


  Query Match             63.5%;  Score 33;  DB 1;  Length 334;
  Best Local Similarity   72.7%;  Pred. No. 98;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLLK 12
           |||| |: |||
Db      86 LLLKALVPLLK 96


RESULT  14
DCAM_MESAU
ID   DCAM_MESAU     STANDARD;      PRT;   334 AA.
AC   P28918;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50) (AdoMetDC)
DE   (SamDC) [Contains: S-adenosylmethionine decarboxylase alpha chain; S-
DE   adenosylmethionine decarboxylase beta chain].
GN   AMD1.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92223099; PubMed=1562599;
RA   Tekwani B.L., Stanley B.A., Pegg A.E.;
RT   "Nucleotide sequence of hamster S-adenosylmethionine decarboxylase
RT   cDNA.";
RL   Biochim. Biophys. Acta 1130:221-223(1992).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = (5-deoxy-5-
CC       adenosyl)(3-aminopropyl) methylsulfonium salt + CO(2).
CC   -!- COFACTOR: PYRUVOYL GROUP.
CC   -!- PATHWAY: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE
CC       AMINOPROPYL MOIETY REQUIRED FOR SPERMIDINE AND SPERMINE
CC       BIOSYNTHESIS FROM PUTRESCINE.
CC   -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC ADOMETDC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X63861; CAA45343.1; -.
DR   PIR; S19871; DCHYDM.
DR   PIR; S22358; S22358.
DR   HSSP; P17707; 1JEN.
DR   InterPro; IPR001985; SAM_decarbox.
DR   Pfam; PF01536; SAM_decarbox; 1.
DR   ProDom; PD002379; SAM_decarbox; 1.
DR   PROSITE; PS01336; ADOMETDC; 1.
KW   Spermidine biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen.
FT   CHAIN         1     67          S-ADENOSYLMETHIONINE DECARBOXYLASE BETA
FT                                   CHAIN.
FT   CHAIN        68    334          S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA
FT                                   CHAIN.
FT   SITE         67     68          CLEAVAGE (NONHYDROLYTIC).
FT   MOD_RES      68     68          CONVERTED TO A PYRUVOYL GROUP.
FT   ACT_SITE      8      8          IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                   SIMILARITY).
FT   ACT_SITE     11     11          IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                   SIMILARITY).
FT   ACT_SITE     82     82          IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                   SIMILARITY).
SQ   SEQUENCE   334 AA;  38313 MW;  FB519BCA749A1A7D CRC64;


  Query Match             63.5%;  Score 33;  DB 1;  Length 334;
  Best Local Similarity   72.7%;  Pred. No. 98;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLLK 12
           |||| |: |||
Db      86 LLLKALVPLLK 96


RESULT  15
DCM1_MOUSE
ID   DCM1_MOUSE     STANDARD;      PRT;   334 AA.
AC   P31154;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   S-adenosylmethionine decarboxylase proenzyme 1 (EC 4.1.1.50) (AdoMetDC
DE   1) (SamdC 1) [Contains: S-adenosylmethionine decarboxylase 1 alpha
DE   chain; S-adenosylmethionine decarboxylase 1 beta chain].
GN   AMD1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=93080592; PubMed=1449493;
RA   Waris T., Ihalainen R., Keraenen M.-R., Pajunen A.;
RT   "Molecular cloning of the mouse S-adenosylmethionine decarboxylase
RT   cDNA: specific protein binding to the conserved region of the mRNA
RT   5'-untranslated region.";
RL   Biochem. Biophys. Res. Commun. 189:424-429(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93345510; PubMed=8344293;
RA   Suzuki T., Sadakata Y., Kashiwagi K., Hoshino K., Kakinuma Y.,
RA   Shirahata A., Igarashi K.;
RT   "Overproduction of S-adenosylmethionine decarboxylase in
RT   ethylglyoxal-bis(guanylhydrazone)-resistant mouse FM3A cells.";
RL   Eur. J. Biochem. 215:247-253(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ; TISSUE=Spleen;
RX   MEDLINE=20035739; PubMed=10570962;
RA   Nishimura K., Kashiwagi K., Matsuda Y., Jaenne O.A., Igarashi K.;
RT   "Gene structure and chromosomal localization of mouse
RT   S-adenosylmethionine decarboxylase.";
RL   Gene 238:343-350(1999).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = (5-deoxy-5-
CC       adenosyl)(3-aminopropyl) methylsulfonium salt + CO(2).
CC   -!- COFACTOR: PYRUVOYL GROUP.
CC   -!- PATHWAY: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE
CC       AMINOPROPYL MOIETY REQUIRED FOR SPERMIDINE AND SPERMINE
```

```
CC       BIOSYNTHESIS FROM PUTRESCINE.
CC   -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE EUKARYOTIC ADOMETDC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z14986; CAA78710.1; -.
DR   EMBL; D12780; BAA02243.1; -.
DR   EMBL; AB025024; BAA83784.1; -.
DR   HSSP; P17707; 1JEN.
DR   MGD; MGI:88004; Amd1.
DR   InterPro; IPR001985; SAM_decarbox.
DR   Pfam; PF01536; SAM_decarbox; 1.
DR   ProDom; PD002379; SAM_decarbox; 1.
DR   PROSITE; PS01336; ADOMETDC; 1.
KW   Spermidine biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen.
FT   CHAIN         1     67       S-ADENOSYLMETHIONINE DECARBOXYLASE 1 BETA
FT                                CHAIN.
FT   CHAIN        68    334       S-ADENOSYLMETHIONINE DECARBOXYLASE 1
FT                                ALPHA CHAIN.
FT   SITE         67     68       CLEAVAGE (NONHYDROLYTIC).
FT   MOD_RES      68     68       CONVERTED TO A PYRUVOYL GROUP.
FT   ACT_SITE      8      8       IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                SIMILARITY).
FT   ACT_SITE     11     11       IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                SIMILARITY).
FT   ACT_SITE     82     82       IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT                                SIMILARITY).
SQ   SEQUENCE   334 AA;  38272 MW;  7950A1E9A9ACBD72 CRC64;


  Query Match             63.5%;  Score 33;  DB 1;  Length 334;
  Best Local Similarity   72.7%;  Pred. No. 98;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLLK 12
           |||| |: |||
Db      86 LLLKALVPLLK 96


Search completed: June 17, 2002, 12:44:45
Job time: 300 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:39:25 ; Search time 73.61 Seconds
(without alignments)
28.202 Million cell updates/sec

Title: US-09-367-714A-23
Perfect score: 52
Sequence: 1 KLLLKLLLKLLK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	76.9	2481	10	Q9FR53	Q9fr53 arabidopsis
2	40	76.9	2513	10	Q9LPM4	Q91pm4 arabidopsis
3	37	71.2	137	16	Q97MS6	Q97ms6 clostridium
4	36	69.2	53	8	Q9BC70	Q9bc70 sargassum p
5	36	69.2	74	10	Q94LQ5	Q94lq5 oryza sativ
6	36	69.2	237	4	Q96LK7	Q961k7 homo sapien
7	36	69.2	1107	5	Q95XU2	Q95xu2 caenorhabdi
8	36	69.2	1453	8	Q9G9H3	Q9g9h3 schizophyll
9	36	69.2	1896	10	O64604	O64604 arabidopsis
10	35	67.3	263	16	Q9CP24	Q9cp24 pasteurella
11	35	67.3	318	16	Q9PI84	Q9pi84 campylobact
12	35	67.3	319	16	O51049	O51049 borrelia bu
13	35	67.3	1230	4	Q9Y650	Q9y650 homo sapien
14	35	67.3	1307	4	Q9HC09	Q9hc09 homo sapien
15	35	67.3	1997	4	Q9HC10	Q9hc10 homo sapien
16	35	67.3	3268	3	Q03280	Q03280 saccharomyc
17	34	65.4	48	5	Q9GR26	Q9gr26 aphidius co
18	34	65.4	53	12	O55758	O55758 chilo iride
19	34	65.4	68	12	Q91FX6	Q91fx6 chilo iride
20	34	65.4	191	17	Q97WI9	Q97wi9 sulfolobus
21	34	65.4	208	16	Q9AC69	Q9ac69 staphylococ
22	34	65.4	213	10	O81431	O81431 arabidopsis
23	34	65.4	282	16	Q9ZLZ4	Q9zlz4 helicobacte
24	34	65.4	289	17	Q96ZX0	Q96zx0 sulfolobus
25	34	65.4	386	17	Q973P4	Q973p4 sulfolobus
26	34	65.4	552	17	O58557	O58557 pyrococcus
27	34	65.4	662	16	Q97RB1	Q97rb1 streptococc
28	34	65.4	707	16	Q9PE08	Q9pe08 xylella fas
29	34	65.4	1234	5	Q9VZ85	Q9vz85 drosophila
30	34	65.4	1300	5	Q9NKD6	Q9nkd6 drosophila
31	34	65.4	1360	5	Q9NDI1	Q9ndi1 drosophila
32	34	65.4	1712	5	O96160	O96160 plasmodium
33	34	65.4	2447	5	Q9NEF9	Q9nef9 drosophila
34	33	63.5	45	12	Q91FK0	Q91fk0 chilo iride
35	33	63.5	84	16	Q99YW0	Q99yw0 streptococc
36	33	63.5	84	16	Q97MQ6	Q97mq6 clostridium
37	33	63.5	91	16	O51404	O51404 borrelia bu
38	33	63.5	100	5	Q9NLV3	Q9nlv3 leishmania
39	33	63.5	100	16	Q932M8	Q932m8 staphylococ
40	33	63.5	144	5	P90741	P90741 caenorhabdi
41	33	63.5	144	8	Q9G905	Q9g905 ochromonas
42	33	63.5	162	8	Q950Q3	Q950q3 spizellomyc
43	33	63.5	187	16	Q97S99	Q97s99 streptococc
44	33	63.5	191	17	Q9YFW9	Q9yfw9 aeropyrum p
45	33	63.5	223	16	Q99RZ4	Q99rz4 staphylococ

ALIGNMENTS

RESULT 1
Q9FR53
ID Q9FR53 PRELIMINARY; PRT; 2481 AA.
AC Q9FR53;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PTOR.
GN TOR.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Menand B., Nussaume L., Meyer C., Desnos T., Bouchez D., Robaglia C.;
RT "Mutation in AtTOR affects embryo development.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF178967; AAG43423.1; -.
DR HSSP; P42345; 1FAP.
DR InterPro; IPR003151; FAT.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000403; PI3_PI4_kinase.
DR Pfam; PF02259; FAT; 1.
DR Pfam; PF02260; FATC; 1.
DR Pfam; PF00454; PI3_PI4_kinase; 1.
DR SMART; SM00146; PI3Kc; 1.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS50290; PI3_4_KINASE_3; 1.
SQ SEQUENCE 2481 AA; 279187 MW; DA663EA9A9366F93 CRC64;

Query Match 76.9%; Score 40; DB 10; Length 2481;
Best Local Similarity 90.9%; Pred. No. 1.7e+02;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LLLKLLLKLLK 12

```
          ||| |||||||
Db    785 LLLGLLLKLLK 795


RESULT   2
Q9LPM4
ID   Q9LPM4      PRELIMINARY;      PRT;  2513 AA.
AC   Q9LPM4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   F2J10.9 PROTEIN.
GN   F2J10.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Sakano H., Liu S.X., Yu G., Lee J., Lenz C., Pham P., Toriumi M.,
RA   Chin C., Chiou J., Choi E., Chung M., Gonzalez A., Howng B., Liu A.,
RA   Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA   Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA   Miranda M., Nguyen M., Palm C., Shinn P., Southwick A., Davis R.W.,
RA   Ecker J.R., Federspiel N.A., Theologis A.;
RT   "The sequence of BAC F2J10 from Arabidopsis thaliana chromosome 1.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC015445; AAF76442.1; -.
DR   HSSP; P42345; 1FAP.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
SQ   SEQUENCE   2513 AA;  282911 MW;  A4B9740321AC5261 CRC64;


  Query Match             76.9%;  Score 40;  DB 10;  Length 2513;
  Best Local Similarity   90.9%;  Pred. No. 1.7e+02;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLLK 12
           ||| |||||||
Db     797 LLLGLLLKLLK 807


RESULT   3
Q97MS6
ID   Q97MS6      PRELIMINARY;      PRT;   137 AA.
AC   Q97MS6;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   UNCHARACTERIZED PROTEIN, YJE/RRF2 FAMILY.
GN   CAC0115.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
DR   EMBL; AE007524; AAK78100.1; -.
DR   InterPro; IPR000944; UPF0074.
DR   Pfam; PF02082; UPF0074; 1.
DR   ProDom; PD003632; UPF0074; 1.
DR   PROSITE; PS01332; UPF0074; 1.
KW   Complete proteome.
SQ   SEQUENCE   137 AA;  15645 MW;  C2400AA028BA5DFE CRC64;


  Query Match             71.2%;  Score 37;  DB 16;  Length 137;
  Best Local Similarity   66.7%;  Pred. No. 44;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           : ||||| ||:|
Db      40 RFLLKLLRKLIK 51


RESULT   4
Q9BC70
ID   Q9BC70      PRELIMINARY;      PRT;    53 AA.
AC   Q9BC70;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO
DE   LARGE SUBUNIT) (FRAGMENT).
GN   RBCL.
OS   Sargassum polyceratium.
OG   Chloroplast.
OC   Eukaryota; stramenopiles; Phaeophyceae; Fucales; Sargassaceae;
OC   Sargassum.
OX   NCBI_TaxID=143167;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Phillips N., Fredericq S.;
RT   "Biogeographic and phylogenetic investigation of the pan-pacific genus
RT   Sargassum (Fucales, Phaeophyceae) with respect to the Gulf of Mexico
RT   species.";
RL   Gulf Mex. Sci. 18:1-11(2000).
CC   -!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-
CC       RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC       CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC       THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC       REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC       ACTIVE SITE (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) = 2 3-
CC       PHOSPHO-D-GLYCERATE.
CC   -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) = 3-
CC       PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC   -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
DR   EMBL; AF301225; AAK01554.1; -.
DR   InterPro; IPR000685; RuBisCO_large.
DR   Pfam; PF00016; RuBisCO_large; 1.
KW   Carbon dioxide fixation; Chloroplast; Lyase; Monooxygenase;
KW   Oxidoreductase; Photorespiration; Photosynthesis.
FT   NON_TER       1      1
SQ   SEQUENCE   53 AA;  6293 MW;  6E4CD0DCA9CE531B CRC64;


  Query Match             69.2%;  Score 36;  DB 8;  Length 53;
  Best Local Similarity   72.7%;  Pred. No. 29;
  Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLLK 12
           :|||| ||:||
Db      25 ILLKLQLKVLK 35
```

```
RESULT   5
Q94LQ5
ID   Q94LQ5      PRELIMINARY;      PRT;    74 AA.
AC   Q94LQ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 8.1 KDA PROTEIN.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA   Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA   Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., VanAken S.E.,
RA   Utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA   White O., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBa0010C11 genomic sequence.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC069300; AAK55450.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   74 AA;  8061 MW;  720EF0A677709444 CRC64;


  Query Match             69.2%;  Score 36;  DB 10;  Length 74;
  Best Local Similarity   66.7%;  Pred. No. 38;
  Matches    8;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           :|||:| |:|||
Db      50 RLLLRLWLQLLK 61


RESULT   6
Q96LK7
ID   Q96LK7      PRELIMINARY;      PRT;   237 AA.
AC   Q96LK7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CDNA FLJ25415 FIS, CLONE TST03443.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA   Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA   Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA   Suzuki Y., Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA   Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA   Kawakami B., Nagai K., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK058144; BAB71684.1; -.
SQ   SEQUENCE   237 AA;  25880 MW;  58250E60EE1D3C5A CRC64;


  Query Match             69.2%;  Score 36;  DB 4;  Length 237;
  Best Local Similarity   75.0%;  Pred. No. 1e+02;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | | | ||||||
Db     105 KSLTKYLLKLLK 116


RESULT   7
Q95XU2
ID   Q95XU2      PRELIMINARY;      PRT;  1107 AA.
AC   Q95XU2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 127.2 KDA PROTEIN.
GN   Y67D8A.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Edwards J., Lamar B., Minx P., Du H., Kemp K., Wohldmann P.,
RA   Walker C.;
RT   "The sequence of C. elegans cosmid Y67D8A.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024848; AAK68543.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   1107 AA;  127230 MW;  A27CD8BAF85A81FC CRC64;


  Query Match             69.2%;  Score 36;  DB 5;  Length 1107;
  Best Local Similarity   70.0%;  Pred. No. 3.9e+02;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLKLLLKLLK 12
           :|||||:|:|
Db     982 ILKLLLRLIK 991


RESULT   8
Q9G9H3
ID   Q9G9H3      PRELIMINARY;      PRT;  1453 AA.
AC   Q9G9H3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RIBOSOMAL PROTEIN S3.
GN   RPS3.
OS   Schizophyllum commune (Bracket fungus).
OG   Mitochondrion.
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC   Stereales; Schizophyllaceae; Schizophyllum.
OX   NCBI_TaxID=5334;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20377911; PubMed=10916154;
RA   Bullerwell C.E., Burger G., Lang B.F.;
RT   "A novel motif for identifying rps3 homologs in fungal mitochondrial
RT   genomes.";
RL   Trends Biochem. Sci. 25:363-365(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RA   Forget L., Ustinova J., Wang Z., Huss V.A.R., Lang F.B.F.;
RT   "Hyaloraphidium curvatum: a linear mitochondrial genome, tRNA editing,
RT   and an evolutionary link to lower fungi.";
RL   Mol. Biol. Evol. 0:0-0(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Lang F.B.F.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF402141; AAG10295.1; -.
KW   Mitochondrion.
SQ   SEQUENCE   1453 AA;  170925 MW;  C860EDB511721651 CRC64;


  Query Match             69.2%;  Score 36;  DB 8;  Length 1453;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches    8;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           |:| |||||:|:
Db     410 KVLNKLLLKILE 421


RESULT   9
O64604
ID   O64604      PRELIMINARY;      PRT;  1896 AA.
AC   O64604;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   F17O7.14 PROTEIN.
GN   F17O7.14.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Kwan A., Oji O.,
RA   Liu S., Li J., Araujo R., Au M., Brendel V., Buehler E., Conway A.B.,
RA   Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y., Palm C.J.,
RA   Shinn P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,
RA   Theologis A.;
RT   "Arabidopsis thaliana chromosome 1 BAC F17O7 sequence.";
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Theologis A.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC003671; AAC18812.1; -.
DR   InterPro; IPR002950; Josephin.
DR   InterPro; IPR000449; UBA.
DR   InterPro; IPR003903; UIM.
DR   Pfam; PF00627; UBA; 1.
DR   Pfam; PF02809; UIM; 1.
DR   SMART; SM00165; UBA; 1.
SQ   SEQUENCE   1896 AA;  210020 MW;  6659881792E52D8A CRC64;


  Query Match             69.2%;  Score 36;  DB 10;  Length 1896;
  Best Local Similarity   72.7%;  Pred. No. 6.2e+02;
  Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLLK 12
           |||:|:|| ||
Db    1832 LLLELMLKTLK 1842


RESULT  10
Q9CP24
ID   Q9CP24      PRELIMINARY;      PRT;  263 AA.
AC   Q9CP24;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   YEBM.
GN   YEBM OR PM0242.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PM70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR   EMBL; AE006058; AAK02326.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
KW   Complete proteome.
SQ   SEQUENCE   263 AA;  29399 MW;  2DA8CAE6EE1D4FCB CRC64;


  Query Match             67.3%;  Score 35;  DB 16;  Length 263;
  Best Local Similarity   88.9%;  Pred. No. 1.7e+02;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLKLLLKLL 11
           |||:|||||
Db      52 LLKVLLKLL 60


RESULT  11
Q9PI84
ID   Q9PI84      PRELIMINARY;      PRT;  318 AA.
AC   Q9PI84;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE INTEGRAL MEMBRANE PROTEIN.
GN   CJ0421C.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC   Campylobacter.
OX   NCBI_TaxID=197;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 11168;
RX   MEDLINE=20150912; PubMed=10688204;
RA   Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA   Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA   Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA   Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA   Whitehead S., Barrell B.G.;
RT   "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT   reveals hypervariable sequences.";
RL   Nature 403:665-668(2000).
DR   EMBL; AL139075; CAB74257.1; -.
KW   Complete proteome.
SQ   SEQUENCE   318 AA;  37371 MW;  E26F5D88241E2968 CRC64;


  Query Match             67.3%;  Score 35;  DB 16;  Length 318;
  Best Local Similarity   58.3%;  Pred. No. 2e+02;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           ||:|| : |:||
Db     233 KLILKQIAKILK 244
```

```
RESULT  12
O51049
ID   O51049      PRELIMINARY;      PRT;   319 AA.
AC   O51049;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CONSERVED HYPOTHETICAL INTEGRAL MEMBRANE PROTEIN.
GN   BB0017.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
DR   EMBL; AE001116; AAC66414.1; -.
DR   TIGR; BB0017; -.
DR   InterPro; IPR003740; DUF161.
DR   Pfam; PF02588; DUF161; 1.
KW   Complete proteome.
SQ   SEQUENCE   319 AA;  35178 MW;  289D8371C6209DBC CRC64;


  Query Match             67.3%;  Score 35;  DB 16;  Length 319;
  Best Local Similarity   66.7%;  Pred. No. 2e+02;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           |||| :|:| ||
Db      17 KLLLLILIKKLK 28


RESULT  13
Q9Y650
ID   Q9Y650      PRELIMINARY;      PRT;  1230 AA.
AC   Q9Y650;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OTOFERLIN.
GN   OTOF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99206603; PubMed=10192385;
RA   Yasunaga S., Grati M., Cohen-Salmon M., El-Amraoui A., Mustapha M.,
RA   Salem N., El-Zir E., Loiselet J., Petit C.;
RT   "A mutation in OTOF, encoding otoferlin, a FER-1-like protein, causes
RT   DFNB9, a nonsyndromic form of deafness.";
RL   Nat. Genet. 21:363-369(1999).
DR   EMBL; AF107403; AAD26117.1; -.
DR   HSSP; P04410; 1A25.
DR   InterPro; IPR000008; C2.
DR   Pfam; PF00168; C2; 2.
DR   PRINTS; PR00360; C2DOMAIN.
DR   SMART; SM00239; C2; 3.
DR   PROSITE; PS00499; C2_DOMAIN_1; 2.
DR   PROSITE; PS50004; C2_DOMAIN_2; 2.
SQ   SEQUENCE   1230 AA;  140496 MW;  2F0CA02F4877AB48 CRC64;


  Query Match             67.3%;  Score 35;  DB 4;  Length 1230;
  Best Local Similarity   90.0%;  Pred. No. 6.3e+02;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLL 11
           ||||||| ||
Db    1198 LLLKLLLLLL 1207


RESULT  14
Q9HC09
ID   Q9HC09      PRELIMINARY;      PRT;  1307 AA.
AC   Q9HC09;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BRAIN OTOFERLIN SHORT ISOFORM.
GN   OTOF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20395831; PubMed=10903124;
RA   Yasunaga S., Grati M., Chardenoux S., Smith T.N., Friedman T.B.,
RA   Lalwani A.K., Wilcox E.R., Petit C.;
RT   "OTOF encodes multiple long and short isoforms: genetic evidence that
RT   the long ones underlie recessive deafness DFNB9.";
RL   Am. J. Hum. Genet. 67:591-600(2000).
DR   EMBL; AF183186; AAG12992.1; -.
DR   HSSP; P04410; 1A25.
DR   InterPro; IPR000008; C2.
DR   Pfam; PF00168; C2; 2.
DR   PRINTS; PR00360; C2DOMAIN.
DR   SMART; SM00239; C2; 3.
DR   PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR   PROSITE; PS50004; C2_DOMAIN_2; 2.
SQ   SEQUENCE   1307 AA;  148926 MW;  CCCF84A6464A5462 CRC64;


  Query Match             67.3%;  Score 35;  DB 4;  Length 1307;
  Best Local Similarity   90.0%;  Pred. No. 6.6e+02;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLL 11
           ||||||| ||
Db    1275 LLLKLLLLLL 1284


RESULT  15
Q9HC10
ID   Q9HC10      PRELIMINARY;      PRT;  1997 AA.
AC   Q9HC10;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BRAIN OTOFERLIN LONG ISOFORM.
GN   OTOF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20395831; PubMed=10903124;
RA   Yasunaga S., Grati M., Chardenoux S., Smith T.N., Friedman T.B.,
RA   Lalwani A.K., Wilcox E.R., Petit C.;
```

```
RT   "OTOF encodes multiple long and short isoforms: genetic evidence that
RT   the long ones underlie recessive deafness DFNB9.";
RL   Am. J. Hum. Genet. 67:591-600(2000).
DR   EMBL; AF183185; AAG12991.1; -.
DR   HSSP; P04410; 1A25.
DR   InterPro; IPR000008; C2.
DR   Pfam; PF00168; C2; 4.
DR   PRINTS; PR00360; C2DOMAIN.
DR   SMART; SM00239; C2; 6.
DR   PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR   PROSITE; PS50004; C2_DOMAIN_2; 4.
SQ   SEQUENCE   1997 AA;  226751 MW;  24DE196371FB7385 CRC64;

  Query Match             67.3%;  Score 35;  DB 4;  Length 1997;
  Best Local Similarity   90.0%;  Pred. No. 9.5e+02;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLLKLLLKLL 11
           ||||||| ||
Db    1965 LLLKLLLLLL 1974


Search completed: June 17, 2002, 12:44:18
Job time: 293 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:38:20 ; Search time 34.71 Seconds
(without alignments)
8.444 Million cell updates/sec

Title: US-09-367-714A-23
Perfect score: 52
Sequence: 1 KLLLKLLLKLLK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	76.9	14	2	US-08-569-188-8	Sequence 8, Appli
2	40	76.9	14	5	PCT-US94-07019-8	Sequence 8, Appli
3	40	76.9	16	2	US-08-569-188-1	Sequence 1, Appli
4	40	76.9	16	2	US-08-569-188-10	Sequence 10, Appl
5	40	76.9	16	2	US-08-569-188-11	Sequence 11, Appl
6	40	76.9	16	2	US-08-569-188-12	Sequence 12, Appl
7	40	76.9	16	5	PCT-US94-07019-1	Sequence 1, Appli
8	40	76.9	16	5	PCT-US94-07019-10	Sequence 10, Appl
9	40	76.9	16	5	PCT-US94-07019-11	Sequence 11, Appl
10	40	76.9	16	5	PCT-US94-07019-12	Sequence 12, Appl
11	40	76.9	17	2	US-08-569-188-3	Sequence 3, Appli
12	40	76.9	17	2	US-08-818-253-39	Sequence 39, Appl
13	40	76.9	17	4	US-08-818-252-39	Sequence 39, Appl
14	40	76.9	17	5	PCT-US94-07019-3	Sequence 3, Appli
15	40	76.9	18	2	US-08-569-188-5	Sequence 5, Appli
16	40	76.9	18	5	PCT-US94-07019-5	Sequence 5, Appli
17	40	76.9	22	1	US-07-725-331-60	Sequence 60, Appl
18	40	76.9	22	5	PCT-US91-05047-60	Sequence 60, Appl
19	40	76.9	23	2	US-08-290-853-19	Sequence 19, Appl
20	40	76.9	26	1	US-07-725-331-61	Sequence 61, Appl
21	40	76.9	26	5	PCT-US91-05047-61	Sequence 61, Appl
22	40	76.9	30	1	US-07-725-331-62	Sequence 62, Appl
23	40	76.9	30	5	PCT-US91-05047-62	Sequence 62, Appl
24	40	76.9	36	1	US-07-725-331-63	Sequence 63, Appl
25	40	76.9	36	5	PCT-US91-05047-63	Sequence 63, Appl
26	40	76.9	40	2	US-08-687-551-6	Sequence 6, Appli
27	38	73.1	21	1	US-08-944-133-13	Sequence 13, Appl
28	37	71.2	14	1	US-07-725-331-1	Sequence 1, Appli
29	37	71.2	14	5	PCT-US91-05047-1	Sequence 1, Appli
30	37	71.2	16	2	US-08-569-188-2	Sequence 2, Appli
31	37	71.2	16	2	US-08-569-188-13	Sequence 13, Appl
32	37	71.2	16	5	PCT-US94-07019-2	Sequence 2, Appli
33	37	71.2	16	5	PCT-US94-07019-13	Sequence 13, Appl
34	37	71.2	17	2	US-08-569-188-4	Sequence 4, Appli
35	37	71.2	17	2	US-08-569-188-14	Sequence 14, Appl
36	37	71.2	17	5	PCT-US94-07019-4	Sequence 4, Appli
37	37	71.2	17	5	PCT-US94-07019-14	Sequence 14, Appl
38	37	71.2	18	2	US-08-569-188-6	Sequence 6, Appli
39	37	71.2	18	2	US-08-569-188-15	Sequence 15, Appl
40	37	71.2	18	4	US-08-960-054A-12	Sequence 12, Appl
41	37	71.2	18	4	US-08-958-993A-12	Sequence 12, Appl
42	37	71.2	18	4	US-09-296-089-36	Sequence 36, Appl
43	37	71.2	18	5	PCT-US94-07019-6	Sequence 6, Appli
44	37	71.2	18	5	PCT-US94-07019-15	Sequence 15, Appl
45	35	67.3	16	1	US-07-725-331-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT   1
US-08-569-188-8
; Sequence 8, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-8


  Query Match             76.9%;  Score 40;  DB 2;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 1.5;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | ||||| ||||
```

```
Db      2 KKLLKLLKKLLK 13

RESULT   2
PCT-US94-07019-8
; Sequence 8, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-8


  Query Match             76.9%;  Score 40;  DB 5;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 1.5;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLLKLLK 12
          | |||||  ||||
Db      2 KKLLKLLKKLLK 13


RESULT   3
US-08-569-188-1
; Sequence 1, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-1


  Query Match             76.9%;  Score 40;  DB 2;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLLKLLK 12
          | |||||  ||||
Db      4 KKLLKLLKKLLK 15


RESULT   4
US-08-569-188-10
; Sequence 10, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-10


  Query Match             76.9%;  Score 40;  DB 2;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLLKLLK 12
          | |||||  ||||
Db      4 KKLLKLLKKLLK 15
```

```
RESULT   5
US-08-569-188-11
; Sequence 11, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-11


  Query Match             76.9%;  Score 40;  DB 2;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15


RESULT   6
US-08-569-188-12
; Sequence 12, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-12


  Query Match             76.9%;  Score 40;  DB 2;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15


RESULT   7
PCT-US94-07019-1
; Sequence 1, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-1


  Query Match             76.9%;  Score 40;  DB 5;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15


RESULT   8
PCT-US94-07019-10
; Sequence 10, Application PC/TUS9407019
```

```
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-10


  Query Match             76.9%;  Score 40;  DB 5;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15


RESULT  9
PCT-US94-07019-11
; Sequence 11, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-11


  Query Match             76.9%;  Score 40;  DB 5;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15


RESULT  10
PCT-US94-07019-12
; Sequence 12, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-12


  Query Match             76.9%;  Score 40;  DB 5;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15


RESULT  11
US-08-569-188-3
; Sequence 3, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-3


  Query Match             76.9%;  Score 40;  DB 2;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 1.8;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | |||||  ||||
Db       5 KKLLKLLKKLLK 16


RESULT  12
US-08-818-253-39
; Sequence 39, Application US/08818253
; Patent No. 5998204
;  GENERAL INFORMATION:
;    APPLICANT:  Tsien, Roger Y.
;    APPLICANT:  Miyawaki, Atsushi
;    TITLE OF INVENTION:  FLUORESCENT PROTEIN SENSORS FOR
;    TITLE OF INVENTION:  DETECTION OF ANALYTES
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,253
;      FILING DATE:  14-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Ph.D., Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07257/043001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-818-253-39


  Query Match             76.9%;  Score 40;  DB 2;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 1.8;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | |||||  ||||
Db       4 KKLLKLLKKLLK 15


RESULT  13
US-08-818-252-39
; Sequence 39, Application US/08818252B
; Patent No. 6197928
; GENERAL INFORMATION:
;  APPLICANT: Tsien, Roger Y.
;  APPLICANT: Miyawaki, Atsushi
;  TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
;  TITLE OF INVENTION: DETECTION OF ANALYTES
;  FILE REFERENCE: 07257/042001
;  CURRENT APPLICATION NUMBER: US/08/818,252B
;  CURRENT FILING DATE: 1997-03-14
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Calmodulin binding peptide-2
US-08-818-252-39


  Query Match             76.9%;  Score 40;  DB 4;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 1.8;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | |||||  ||||
Db       4 KKLLKLLKKLLK 15


RESULT  14
PCT-US94-07019-3
; Sequence 3, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-3


  Query Match             76.9%;  Score 40;  DB 5;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 1.8;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | |||||  ||||
Db       5 KKLLKLLKKLLK 16


RESULT  15
US-08-569-188-5
; Sequence 5, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;       STREET:  1007 MARKET STREET
;       CITY:  WILMINGTON
;       STATE:  DELAWARE
;       COUNTRY:  UNITED STATES OF AMERICA
;       ZIP:  19898
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  DISKETTE, 3.50 INCH
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;       SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/569,188
;       FILING DATE:
;       CLASSIFICATION:  525
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/082,852
;       FILING DATE:  JUNE 22, 1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  LINDA AXAMETHY FLOYD
;       REGISTRATION NUMBER:  33,692
;       REFERENCE/DOCKET NUMBER:  CR-9295-A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  302-892-8112
;       TELEFAX:  302-773-0164
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  18 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  peptide
US-08-569-188-5


  Query Match             76.9%;  Score 40;  DB 2;  Length 18;
  Best Local Similarity   83.3%;  Pred. No. 1.9;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLLKLLK 12
           | ||||| ||||
Db       6 KKLLKLLKKLLK 17


Search completed: June 17, 2002, 12:42:03
Job time: 223 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:41:21 ; Search time 94.14 Seconds
(without alignments)
7.079 Million cell updates/sec

Title: US-09-367-714A-28
Perfect score: 26
Sequence: 1 KLLLLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	100.0	6	18	AAW35154	Leu/Lys diastereom
2	26	100.0	6	19	AAW82852	Antipathogenic pep
3	26	100.0	6	21	AAB17418	Antipathogenic pep
4	26	100.0	8	19	AAW45766	KL-4 pulmonary sur
5	26	100.0	8	19	AAW45768	KL-4 pulmonary sur
6	26	100.0	12	18	AAW35166	Leu/Lys diastereom
7	26	100.0	12	19	AAW82883	Antipathogenic pep
8	26	100.0	12	21	AAB17412	Antipathogenic pep
9	26	100.0	13	19	AAW45767	KL-4 pulmonary sur
10	26	100.0	15	19	AAW77384	Lytic peptide with
11	26	100.0	18	19	AAW45765	KL-4 pulmonary sur
12	26	100.0	21	14	AAR30661	Peptide contg. alt
13	26	100.0	21	18	AAW32109	Liposomal pulmonar
14	26	100.0	21	19	AAW45763	KL-4 pulmonary sur
15	26	100.0	21	20	AAW82278	Surfactant peptide
16	26	100.0	21	21	AAB06344	Peptide used in th
17	26	100.0	21	21	AAY88039	Respiratory distre
18	26	100.0	21	22	AAB49683	Modified surfactan
19	26	100.0	21	22	AAB46100	Human surfactant p
20	26	100.0	34	21	AAB34780	Human secreted pro
21	26	100.0	35	21	AAB08290	Artificial Surfact
22	26	100.0	39	21	AAB28015	Human secreted pro
23	26	100.0	64	19	AAW85466	Secreted protein e
24	26	100.0	64	21	AAY87304	Human signal pepti
25	26	100.0	121	19	AAW80409	A secreted protein
26	26	100.0	125	22	ABG19410	Novel human diagno
27	26	100.0	133	22	AAO00228	Human polypeptide
28	26	100.0	136	22	AAO13335	Human polypeptide
29	26	100.0	171	20	AAY37473	Amino acid sequenc
30	26	100.0	197	21	AAG55601	Arabidopsis thalia
31	26	100.0	258	21	AAG55600	Arabidopsis thalia
32	26	100.0	259	21	AAG55599	Arabidopsis thalia
33	26	100.0	307	21	AAG06829	Arabidopsis thalia
34	26	100.0	311	22	AAB80628	Environmental stre
35	26	100.0	388	21	AAG46587	Arabidopsis thalia
36	26	100.0	2431	13	AAR25138	SFV4 non-structura
37	24	92.3	16	17	AAW07285	Amphiphilic antimi
38	24	92.3	43	22	AAU17726	Novel human respir
39	24	92.3	66	18	AAW28160	Amino acid sequenc
40	24	92.3	80	21	AAB24480	Human secreted pro
41	24	92.3	145	22	AAU23303	Novel human enzyme
42	24	92.3	145	22	AAU21748	Novel human neopla
43	24	92.3	187	21	AAY66677	Membrane-bound pro
44	24	92.3	187	22	AAU29236	Human PRO polypept
45	24	92.3	187	22	AAB85575	Human glutathione

ALIGNMENTS

RESULT 1
AAW35154
ID AAW35154 standard; peptide; 6 AA.
XX
AC AAW35154;
XX
DT 14-APR-1998 (first entry)
XX
DE Leu/Lys diastereomer peptide [D]-L1,3-K2L4.
XX
KW Leu/Lys diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW Honey bee venom; pardaxin; cytolytic activity; cancer;
KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW agricultural pesticide; cell wall lysis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "D-form residue"
FT Misc-difference 4
FT /note= "D-form residue"
XX
PN WO9731019-A2.
XX
PD 28-AUG-1997.
XX
PF 20-FEB-1997; 97WO-IL00066.
XX
PR 22-FEB-1996; 96IL-0117223.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;

XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer
XX
PS Claim 21; Page 40; 80pp; English.
XX
CC This sequence represents a Leu/Lys diastereomer peptide of the
CC invention. The peptides of the invention have: (a) cytolytic activity on
CC pathogenic cells (pathogens and malignant cells not naturally present in
CC the body); but (b) no haemolytic activity, or such activity only at a
CC concentration significantly higher than that at which they lyse
CC pathogens. The peptides, their complexes and mixtures are used to treat
CC infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC or cancer, in human and veterinary medicine. Also, they can be used as
CC preservatives for food, cosmetics and agricultural produce, or as
CC agricultural pesticides. The absence of haemolytic activity (associated
CC with disturbance of alpha-helical structures) means that the peptides
CC have few if any toxic effects, and those that include D-aa will have
CC increased resistance to proteolytic degradation. Non-haemolytic,
CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC activity, allowing selection of agents for particular applications. Since
CC these random copolymers induce total lysis of bacterial cell walls,
CC resistance to them is unlikely to develop.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 26; DB 18; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 1 kllllk 6

RESULT 2
AAW82852
ID AAW82852 standard; peptide; 6 AA.
XX
AC AAW82852;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 13; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 26; DB 19; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 1 kllllk 6

RESULT 3
AAB17418
ID AAB17418 standard; Peptide; 6 AA.
XX
AC AAB17418;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:522.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX

```
PS   Claim 39; Page 379; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b,  where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC   -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC   where P1, P2, P3, and P4 = are each independently sequences of
CC   pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC   independently linkers; and a, b, c, d, e, and f = are each independently
CC   0 or 1, provided that at least 1 of a and b is 1. The composition can
CC   have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC   activities. DNAs, vectors and host cells from the present invention can
CC   be used for producing pharmaceutical compositions. The compositions are
CC   useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC   The use of an Fc domain (rather than a Fab domain) can provide a longer
CC   half-life or incorporate functions such as Fc receptor binding, protein
CC   A binding, complement fixation, and possibly placental transfer. AAA69443
CC   to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   6 AA;


  Query Match             100.0%;  Score 26;  DB 21;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db      1 kllllk 6


RESULT   4
AAW45766
ID   AAW45766 standard; peptide; 8 AA.
XX
AC   AAW45766;
XX
DT   19-JUN-1998  (first entry)
XX
DE   KL-4 pulmonary surfactant protein precursor peptide.
XX
KW   Liquid phase peptide synthesis; KL-4 pulmonary surfactant protein;
KW   coupling; respiratory distress syndrome; saponification.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminally modified by t-butyloxycarbonyl
FT                          group. Side chain amino group protected by
FT                          benzyloxycarbonyl group"
FT   Modified-site   6
FT                   /note= "Side chain amino group of Lys6 protected by
FT                          benzyloxycarboxyl group"
FT   Modified-site   8
FT                   /note= "Leu-OR, where R is 1-8C alkyl or phenyl
FT                          1-8C alkyl"
XX
PN   WO9802461-A2.
XX
PD   22-JAN-1998.
XX
PF   11-JUL-1997;   97WO-US12163.
XX
PR   17-JUL-1996;   96US-0021455.
XX
PA   (ORTH ) ORTHO PHARM CORP.
XX
PI   Abdel-magid AF,  Eggmann U,  Maryanoff CA,  Thaler A;
PI   Villani FJ;
XX
DR   WPI; 1998-110531/10.
XX
PT   Preparation of KL-4 pulmonary surfactant - using liquid phase
PT   peptide synthesis procedures by coupling appropriate peptide
PT   fragments
XX
PS   Claim 11; Page 27; 30pp; English.
XX
CC   This peptide is used in a novel process for the synthesis of a KL-4
CC   pulmonary surfactant protein. The process comprises: (a) reacting a
CC   3-amino acid peptide residue of formula H-Lys(Z)-Leu-Leu-OH with a 5-
CC   amino acid peptide residue Boc-Leu-Leu-Lys(Z)-Leu-Leu-OH (II) to yield
CC   an 8-amino acid peptide residue, which is successively reacted with the
CC   5-amino acid peptide to form an 18-amino acid peptide of formula
CC   H-Leu-Leu(Lys(Z)-Leu4)3Lys(Z)-OBzl (IIIa); (b) reacting the 18 amino acid
CC   peptide with a 3-residue amino acid peptide of formula H-Leu-Leu-
CC   Lys(Z)-OBzl (X) to form the protected 21-amino acid KL-4 protein; and
CC   (c) removing the protecting group of the 21-amino acid KL-4 protein by
CC   reaction with a suitable acid to form the final KL-4 protein. The
CC   methods can be used for the preparation of the polypeptide component of
CC   the synthetic pulmonary surfactant KL-4 which can be used in the
CC   treatment of respiratory distress syndrome. The saponification process
CC   can provide for the deprotection of a peptide ester protected carboxyl
CC   group with reduced racemisation. The liquid phase peptide processes
CC   provide advantages in solubility and control over unwanted by-products.
XX
SQ   Sequence   8 AA;


  Query Match             100.0%;  Score 26;  DB 19;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db      1 kllllk 6


RESULT   5
AAW45768
ID   AAW45768 standard; peptide; 8 AA.
XX
AC   AAW45768;
XX
DT   19-JUN-1998  (first entry)
XX
DE   KL-4 pulmonary surfactant protein precursor peptide.
XX
KW   Liquid phase peptide synthesis; KL-4 pulmonary surfactant protein;
KW   coupling; respiratory distress syndrome; saponification.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminally modified by t-butyloxycarbonyl
FT                          group"
FT   Modified-site   3
FT                   /note= "Side chain amino group of Lys3 protected by
FT                          benzyloxycarboxyl group"
FT   Modified-site   8
FT                   /note= "Side chain amino group of Lys8 protected by
FT                          benzyloxycarboxyl group. C-terminally
FT                          modifed by OBzl"
XX
PN   WO9802461-A2.
XX
PD   22-JAN-1998.
XX
PF   11-JUL-1997;   97WO-US12163.
XX
PR   17-JUL-1996;   96US-0021455.
XX
```

```
PA   (ORTH ) ORTHO PHARM CORP.
XX
PI   Abdel-magid AF,  Eggmann U,  Maryanoff CA,  Thaler A;
PI   Villani FJ;
XX
DR   WPI; 1998-110531/10.
XX
PT   Preparation of KL-4 pulmonary surfactant - using liquid phase
PT   peptide synthesis procedures by coupling appropriate peptide
PT   fragments
XX
PS   Claim 2a; Page 25; 30pp; English.
XX
CC   This peptide is used in a novel process for the synthesis of a KL-4
CC   pulmonary surfactant protein. The process comprises: (a) reacting a
CC   3-amino acid peptide residue of formula H-Lys(Z)-Leu-Leu-OH with a 5-
CC   amino acid peptide residue Boc-Leu-Leu-Lys(Z)-Leu-Leu-OH (II) to yield
CC   an 8-amino acid peptide residue, which is successively reacted with the
CC   5-amino acid peptide to form an 18-amino acid peptide of formula
CC   H-Leu-Leu(Lys(Z)-Leu4)3Lys(Z)-OBzl (IIIa); (b) reacting the 18 amino acid
CC   peptide with a 3-residue amino acid peptide of formula H-Leu-Leu-
CC   Lys(Z)-OBzl (X) to form the protected 21-amino acid KL-4 protein; and
CC   (c) removing the protecting group of the 21-amino acid KL-4 protein by
CC   reaction with a suitable acid to form the final KL-4 protein. The
CC   methods can be used for the preparation of the polypeptide component of
CC   the synthetic pulmonary surfactant KL-4 which can be used in the
CC   treatment of respiratory distress syndrome. The saponification process
CC   can provide for the deprotection of a peptide ester protected carboxyl
CC   group with reduced racemisation. The liquid phase peptide processes
CC   provide advantages in solubility and control over unwanted by-products.
XX
SQ   Sequence   8 AA;
```

```
  Query Match             100.0%;  Score 26;  DB 19;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db       3 kllllk 8
```

```
RESULT   6
AAW35166
ID   AAW35166 standard; peptide; 12 AA.
XX
AC   AAW35166;
XX
DT   14-APR-1998  (first entry)
XX
DE   Leu/Lys diastereomer peptide [D]-L3,4,8,10-K3L9.
XX
KW   Leu/Lys diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW   Honey bee venom; pardaxin; cytolytic activity; cancer;
KW   non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW   agricultural pesticide; cell wall lysis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 3
FT                   /note= "D-form residue"
FT   Misc-difference 4
FT                   /note= "D-form residue"
FT   Misc-difference 8
FT                   /note= "D-form residue"
FT   Misc-difference 10
FT                   /note= "D-form residue"
FT   Modified-site   12
FT                   /note= "C-terminal amide"
XX
PN   WO9731019-A2.
XX
PD   28-AUG-1997.
XX
PF   20-FEB-1997;   97WO-IL00066.
XX
PR   22-FEB-1996;   96IL-0117223.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Oren Z,  Shai Y;
XX
DR   WPI; 1997-435088/40.
XX
PT   Peptide(s) having selective cytolytic activity - against pathogens
PT   and malignant cells, but no haemolytic activity, used for treating
PT   infections and cancer
XX
PS   Example 3; Page 39; 80pp; English.
XX
CC   This sequence represents a Leu/Lys diastereomer peptide of the
CC   invention. The peptides of the invention have: (a) cytolytic activity on
CC   pathogenic cells (pathogens and malignant cells not naturally present in
CC   the body); but (b) no haemolytic activity, or such activity only at a
CC   concentration significantly higher than that at which they lyse
CC   pathogens. The peptides, their complexes and mixtures are used to treat
CC   infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC   or cancer, in human and veterinary medicine. Also, they can be used as
CC   preservatives for food, cosmetics and agricultural produce, or as
CC   agricultural pesticides. The absence of haemolytic activity (associated
CC   with disturbance of alpha-helical structures) means that the peptides
CC   have few if any toxic effects, and those that include D-aa will have
CC   increased resistance to proteolytic degradation. Non-haemolytic,
CC   cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC   activity, allowing selection of agents for particular applications. Since
CC   these random copolymers induce total lysis of bacterial cell walls,
CC   resistance to them is unlikely to develop.
XX
SQ   Sequence   12 AA;
```

```
  Query Match             100.0%;  Score 26;  DB 18;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db       7 kllllk 12
```

```
RESULT   7
AAW82883
ID   AAW82883 standard; peptide; 12 AA.
XX
AC   AAW82883;
XX
DT   19-MAY-1999  (first entry)
XX
DE   Antipathogenic peptide.
XX
KW   Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW   cancer; infection; disinfectant; contact lens wetting solution;
KW   preservative; pesticide; fungicide; bactericide.
XX
OS   Synthetic.
XX
PN   WO9837090-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;   98WO-IL00081.
XX
PR   20-FEB-1997;   97WO-IL00066.
XX
```

PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Example 3; Page 38; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 26; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 7 kllllk 12

RESULT 8
AAB17412
ID AAB17412 standard; Peptide; 12 AA.
XX
AC AAB17412;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:516.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX

PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 377; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 26; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 7 kllllk 12

RESULT 9
AAW45767
ID AAW45767 standard; peptide; 13 AA.
XX
AC AAW45767;
XX
DT 19-JUN-1998 (first entry)
XX
DE KL-4 pulmonary surfactant protein precursor peptide.
XX
KW Liquid phase peptide synthesis; KL-4 pulmonary surfactant protein;
KW coupling; respiratory distress syndrome; saponification.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminally modified by t-butyloxycarbonyl
FT group"
FT Modified-site 3
FT /note= "Side chain amino group of Lys3 protected by
FT benzyloxycarboxyl group"
FT Modified-site 8
FT /note= "Side chain amino group of Lys8 protected by
FT benzyloxycarboxyl group"
FT Modified-site 13
FT /note= "Side chain amino group of Lys13 protected by
FT benzyloxycarboxyl group. C-terminally
FT modifed by OBzl"
XX

```
PN   WO9802461-A2.
XX
PD   22-JAN-1998.
XX
PF   11-JUL-1997;   97WO-US12163.
XX
PR   17-JUL-1996;   96US-0021455.
XX
PA   (ORTH ) ORTHO PHARM CORP.
XX
PI   Abdel-magid AF,  Eggmann U,  Maryanoff CA,  Thaler A;
PI   Villani FJ;
XX
DR   WPI; 1998-110531/10.
XX
PT   Preparation of KL-4 pulmonary surfactant - using liquid phase
PT   peptide synthesis procedures by coupling appropriate peptide
PT   fragments
XX
PS   Claim 2c; Page 25; 30pp; English.
XX
CC   This peptide is used in a novel process for the synthesis of a KL-4
CC   pulmonary surfactant protein. The process comprises: (a) reacting a
CC   3-amino acid peptide residue of formula H-Lys(Z)-Leu-Leu-OH with a 5-
CC   amino acid peptide residue Boc-Leu-Leu-Lys(Z)-Leu-Leu-OH (II) to yield
CC   an 8-amino acid peptide residue, which is successively reacted with the
CC   5-amino acid peptide to form an 18-amino acid peptide of formula
CC   H-Leu-Leu(Lys(Z)-Leu4)3Lys(Z)-OBzl (IIIa); (b) reacting the 18 amino acid
CC   peptide with a 3-residue amino acid peptide of formula H-Leu-Leu-
CC   Lys(Z)-OBzl (X) to form the protected 21-amino acid KL-4 protein; and
CC   (c) removing the protecting group of the 21-amino acid KL-4 protein by
CC   reaction with a suitable acid to form the final KL-4 protein. The
CC   methods can be used for the preparation of the polypeptide component of
CC   the synthetic pulmonary surfactant KL-4 which can be used in the
CC   treatment of respiratory distress syndrome. The saponification process
CC   can provide for the deprotection of a peptide ester protected carboxyl
CC   group with reduced racemisation. The liquid phase peptide processes
CC   provide advantages in solubility and control over unwanted by-products.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 26;  DB 19;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db       3 kllllk 8


RESULT  10
AAW77384
ID   AAW77384 standard; peptide; 15 AA.
XX
AC   AAW77384;
XX
DT   14-DEC-1998  (first entry)
XX
DE   Lytic peptide with alterable function 9.
XX
KW   Biologically active peptide; hormone; drug; toxin;
KW   lipid bilayer membrane; microorganism; parasite; virus.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Optional N-terminal myristyl"
FT   Modified-site   15
FT                   /note= "C-terminal amide"
XX
PN   WO9841535-A2.
XX
PD   24-SEP-1998.
XX
PF   18-MAR-1998;   98WO-GB00799.
XX
PR   18-MAR-1997;   97GB-0005519.
XX
PA   (ANMA-) ANMAT TECHNOLOGY LTD.
XX
PI   Ajoula HS,  Clarke DJ;
XX
DR   WPI; 1998-521161/44.
XX
PT   New modified peptide(s) - obtained by substitution with an amino
PT   acid which is modifiable by a reaction and replacing other amino
PT   acids which are not to be modified
XX
PS   Example 1; Page 13; 33pp; English.
XX
CC   The peptides AAW77376-W77390 can be modified by the method of the
CC   invention by substituting at least one amino acid of the peptide to
CC   provide a peptide having at least one amino acid which is modifiable by
CC   a reaction and replacing other amino acids in the peptide with amino
CC   acids which are not modifiable by the reaction.  The methods can be used
CC   for the modification of biologically active peptides such as hormones,
CC   drugs, toxins and peptides which act on lipid bilayer membranes.  The
CC   modified peptides can be used e.g. in the body of an animal or plant or
CC   parts in order to affect the structure or integrity or permeability of a
CC   foreign body such as a microorganism, parasite or virus present in the
CC   body of the animal or plant or within the cells of the body of the animal
CC   or plant.
XX
SQ   Sequence   15 AA;

  Query Match             100.0%;  Score 26;  DB 19;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db       7 kllllk 12


RESULT  11
AAW45765
ID   AAW45765 standard; peptide; 18 AA.
XX
AC   AAW45765;
XX
DT   19-JUN-1998  (first entry)
XX
DE   KL-4 pulmonary surfactant protein precursor peptide.
XX
KW   Liquid phase peptide synthesis; KL-4 pulmonary surfactant protein;
KW   coupling; respiratory distress syndrome; saponification.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   3
FT                   /note= "Side chain amino group of Lys protected by
FT                    benzyloxycarboxyl group"
FT   Modified-site   8
FT                   /note= "Side chain amino group of Lys protected by
FT                    benzyloxycarboxyl group"
FT   Modified-site   13
FT                   /note= "Side chain amino group of Lys protected by
FT                    benzyloxycarboxyl group"
FT   Modified-site   18
FT                   /note= "Side chain amino group of Lys protected by
FT                    benzyloxycarboxyl group and C-terminus
FT                    modifed by OBzl"
```

XX
PN WO9802461-A2.
XX
PD 22-JAN-1998.
XX
PF 11-JUL-1997; 97WO-US12163.
XX
PR 17-JUL-1996; 96US-0021455.
XX
PA (ORTH) ORTHO PHARM CORP.
XX
PI Abdel-magid AF, Eggmann U, Maryanoff CA, Thaler A;
PI Villani FJ;
XX
DR WPI; 1998-110531/10.
XX
PT Preparation of KL-4 pulmonary surfactant - using liquid phase
PT peptide synthesis procedures by coupling appropriate peptide
PT fragments
XX
PS Claim 1a; Page 25; 30pp; English.
XX
CC This peptide is used in a novel process for the synthesis of a KL-4
CC pulmonary surfactant protein. The process comprises: (a) reacting a
CC 3-amino acid peptide residue of formula H-Lys(Z)-Leu-Leu-OH with a 5-
CC amino acid peptide residue Boc-Leu-Leu-Lys(Z)-Leu-Leu-OH (II) to yield
CC an 8-amino acid peptide residue, which is successively reacted with the
CC 5-amino acid peptide to form an 18-amino acid peptide of formula
CC H-Leu-Leu(Lys(Z)-Leu4)3Lys(Z)-OBzl (IIIa); (b) reacting the 18 amino acid
CC peptide with a 3-residue amino acid peptide of formula H-Leu-Leu-
CC Lys(Z)-OBzl (X) to form the protected 21-amino acid KL-4 protein; and
CC (c) removing the protecting group of the 21-amino acid KL-4 protein by
CC reaction with a suitable acid to form the final KL-4 protein. The
CC methods can be used for the preparation of the polypeptide component of
CC the synthetic pulmonary surfactant KL-4 which can be used in the
CC treatment of respiratory distress syndrome. The saponification process
CC can provide for the deprotection of a peptide ester protected carboxyl
CC group with reduced racemisation. The liquid phase peptide processes
CC provide advantages in solubility and control over unwanted by-products.
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 26; DB 19; Length 18;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 3 kllllk 8

RESULT 12
AAR30661
ID AAR30661 standard; peptide; 21 AA.
XX
AC AAR30661;
XX
DT 13-MAY-1993 (first entry)
XX
DE Peptide contg. alternating hydrophobic and hydrophilic regions.
XX
KW Pulmonary surfactant; phospholipid; respiratory distress syndrome;
KW RDS.
XX
OS Synthetic.
XX
PN WO9222315-A.
XX
PD 23-DEC-1992.
XX
PF 01-JUN-1992; 92WO-US04537.
XX
PR 14-JUN-1991; 91US-0715397.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Cochrane CG, Revak SD;
XX
DR WPI; 1993-017902/02.
XX
PT Polypeptide(s) comprising alternating hydrophobic and hydrophilic
PT residue regions - are useful as pulmonary surfactants for
PT treating respiratory distress syndrome
XX
PS Claim 4; Page 58; 73pp; English.
XX
CC The peptide is an example of a highly generic peptide comprising 10-
CC 60 amino acid residues, including a sequence having alternating
CC hydrophobic and hydrophilic amino acid residue regions. When the
CC polypeptide is mixed with a phospholipid a synthetic pulmonary
CC surfactant is formed which has greater surfactant activity than the
CC phospholipid alone. The surfactant is used for the treatment of
CC respiratory distress syndrome (RDS).
CC See also AAR30655-64.
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 26; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 6 kllllk 11

RESULT 13
AAW32109
ID AAW32109 standard; peptide; 21 AA.
XX
AC AAW32109;
XX
DT 04-FEB-1998 (first entry)
XX
DE Liposomal pulmonary surfactant composition polypeptide.
XX
KW Liposomal pulmonary surfactant composition; preparation;
KW respiratory distress syndrome; premature infant; premature neonate;
KW monolayer formation; alveolar air-water interface.
XX
OS Synthetic.
XX
PN WO9719108-A1.
XX
PD 29-MAY-1997.
XX
PF 22-OCT-1996; 96WO-US16804.
XX
PR 20-NOV-1995; 95US-0007347.
XX
PA (ORTH) ORTHO PHARM CORP.
XX
PI Kasulanis CF, Sampino K, Weber JV;
XX
DR WPI; 1997-298058/27.
XX
PT Improved preparation of liposomal pulmonary surfactant - for
PT treatment of respiratory distress syndrome in premature infants and
PT neonate(s)
XX
PS Claim 3; Page 16; 31pp; English.
XX
CC The preparation of liposomal pulmonary surfactant composition by
CC ethanolic injection has been improved. The liposomal pulmonary

CC surfactant composition comprises a polypeptide and a phospholipid.
CC The present sequence represents the polypeptide for use in the
CC liposomal pulmonary surfactant composition. The preparation comprises
CC preparing a form of the polypeptide (or a salt or ester), which
CC exhibits enhanced solubility in ethanol, which comprises: (a) preparing
CC a solution of the polypeptide, salt or ester in a fluorinated alcohol
CC at 5-40 mg/ml; (b) incubating for time sufficient to achieve an optical
CC density at 450 nm of < 0.06; (c) filtering; and (d) removing the
CC fluorinated alcohol to recover solid, soluble polypeptide. The
CC liposomal pulmonary surfactant composition can be used to promote the
CC formation of a monolayer at the alveolar air-water interface, and by
CC reducing the surface tension, prevent the collapse of the alveolus
CC during expiration. The liposomal pulmonary surfactant composition can
CC be used in premature infants and occasionally full term neonates who
CC sometimes suffer from respiratory distress syndrome due to the lack of
CC sufficient endogenous liposomal pulmonary surfactant.
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 26; DB 18; Length 21;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 1 kllllk 6

RESULT 14
AAW45763
ID AAW45763 standard; peptide; 21 AA.
XX
AC AAW45763;
XX
DT 19-JUN-1998 (first entry)
XX
DE KL-4 pulmonary surfactant protein.
XX
KW Liquid phase peptide synthesis; KL-4 pulmonary surfactant protein;
KW coupling; respiratory distress syndrome; saponification.
XX
OS Synthetic.
XX
PN WO9802461-A2.
XX
PD 22-JAN-1998.
XX
PF 11-JUL-1997; 97WO-US12163.
XX
PR 17-JUL-1996; 96US-0021455.
XX
PA (ORTH) ORTHO PHARM CORP.
XX
PI Abdel-magid AF, Eggmann U, Maryanoff CA, Thaler A;
PI Villani FJ;
XX
DR WPI; 1998-110531/10.
XX
PT Preparation of KL-4 pulmonary surfactant - using liquid phase
PT peptide synthesis procedures by coupling appropriate peptide
PT fragments
XX
PS Claim 1a; Page 25; 30pp; English.
XX
CC This sequence represents a synthetic KL-4 pulmonary surfactant protein.
CC A novel process for its synthesis comprises: (a) reacting a 3-amino
CC acid peptide residue of formula H-Lys(Z)-Leu-Leu-OH with a 5-amino
CC acid peptide residue Boc-Leu-Leu-Lys(Z)-Leu-Leu-OH (II) to yield
CC an 8-amino acid peptide residue, which is successively reacted with the
CC 5-amino acid peptide to form an 18-amino acid peptide of formula
CC H-Leu-Leu(Lys(Z)-Leu4)3Lys(Z)-OBzl (IIIa); (b) reacting the 18 amino acid
CC peptide with a 3-residue amino acid peptide of formula H-Leu-Leu-
CC Lys(Z)-OBzl (X) to form the protected 21-amino acid KL-4 protein; and
CC (c) removing the protecting group of the 21-amino acid KL-4 protein by
CC reaction with a suitable acid to form the final KL-4 protein. The
CC methods can be used for the preparation of the polypeptide component of
CC the synthetic pulmonary surfactant KL-4 which can be used in the
CC treatment of respiratory distress syndrome. The saponification process
CC can provide for the deprotection of a peptide ester protected carboxyl
CC group with reduced racemisation. The liquid phase peptide processes
CC provide advantages in solubility and control over unwanted by-products.
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 26; DB 19; Length 21;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 1 kllllk 6

RESULT 15
AAW82278
ID AAW82278 standard; Peptide; 21 AA.
XX
AC AAW82278;
XX
DT 15-MAR-1999 (first entry)
XX
DE Surfactant peptide KL4.
XX
KW Surfactant; pulmonary lavage; inflammation; acute hypoaemia;
KW diaphramatic hernia; respiratory distress syndrome;
KW meconium aspiration syndrome; pneumonia; therapy.
XX
OS Synthetic.
XX
PN WO9849191-A1.
XX
PD 05-NOV-1998.
XX
PF 29-JAN-1998; 98WO-US01711.
XX
PR 28-APR-1997; 97US-0848580.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Cochrane CG, Revak SD;
XX
DR WPI; 1999-034654/03.
XX
PT Pulmonary lavage with dilute surfactant solution at positive
PT end-expiratory pressure - with removal of fluid using short periods
PT of suction, used to treat respiratory distress, e.g. in neonates
PT where caused by aspiration of meconium
XX
PS Claim 44; Page 123; 145pp; English.
XX
CC KL4 is a synthetic peptide that can be used in a synthetic
CC pulmonary surfactant of the invention. It is a mimic of human
CC surfactant protein SP-18. Synthetic pulmonary surfactants comprise
CC one or more phospholipids and a peptide having alternating
CC hydrophobic and hydrophilic regions, or alternating groups of
CC charged and uncharged amino acids (see AAW82278-82 and AAW88193-97).
CC The synthetic pulmonary surfactant is used in a claimed method for
CC pulmonary lavage of a mammal. Lavage is used to treat respiratory
CC distress syndrome caused by aspiration of meconium or gastric
CC contents, pulmonary inflammation or infection, acute hypoxaemia,
CC persistent foetal circulation, congenital diaphramatic hernia,
CC sepsis, trauma, pancreatitis, inhalation of hot or noxious vapour,
CC pneumonia or multiple transfusions. The lavage solution removes
CC inflammatory mediators and preserves or restores pulmonary

CC function.
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 26; DB 20; Length 21;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 KLLLLK 6
          ||||||
Db      1 kllllk 6
```

Search completed: June 17, 2002, 12:41:22
Job time: 297 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:42:57 ; Search time 46.42 Seconds
(without alignments)
12.420 Million cell updates/sec

Title: US-09-367-714A-28
Perfect score: 26
Sequence: 1 KLLLLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	100.0	110	2	A47154	conserved hypothet
2	26	100.0	110	2	F81868	probable proline-r
3	26	100.0	126	2	B72621	hypothetical prote
4	26	100.0	205	2	C87309	peptidyl-tRNA hydr
5	26	100.0	277	2	AB2461	hypothetical prote
6	26	100.0	305	2	A70212	conserved hypothet
7	26	100.0	321	2	S51395	hypothetical prote
8	26	100.0	392	2	A81745	conserved hypothet
9	26	100.0	403	2	B71482	hypothetical prote
10	26	100.0	542	2	B81910	probable ABC-trans
11	26	100.0	542	2	E81105	ABC transporter, A
12	26	100.0	608	2	T18437	hypothetical prote
13	26	100.0	641	2	A39961	choline O-acetyltr
14	26	100.0	688	2	S32961	hypothetical prote
15	26	100.0	748	2	A60202	choline O-acetyltr
16	26	100.0	1855	2	S41649	DNA polymerase - m
17	26	100.0	2431	1	MNWVSF	nonstructural poly
18	24	92.3	41	2	T07329	hypothetical prote
19	24	92.3	80	2	H84974	hypothetical prote
20	24	92.3	93	2	S15948	hypothetical prote
21	24	92.3	95	2	E95270	hypothetical prote
22	24	92.3	146	2	C81036	probable marR-fami
23	24	92.3	165	2	A64227	hypothetical prote
24	24	92.3	201	2	AG1668	hypothetical prote
25	24	92.3	215	2	A81693	conserved hypothet
26	24	92.3	215	2	G71537	hypothetical prote
27	24	92.3	217	1	B64600	NAD(P)H-flavin oxi
28	24	92.3	217	2	H71912	probable oxidoredu
29	24	92.3	239	2	T41951	hypothetical prote
30	24	92.3	239	2	A81301	probable anion-upt
31	24	92.3	244	2	T30358	hypothetical prote
32	24	92.3	260	2	E90010	hypothetical prote
33	24	92.3	262	2	A99155	conserved hypothet
34	24	92.3	294	2	T19055	hypothetical prote
35	24	92.3	325	2	T31977	hypothetical prote
36	24	92.3	328	2	H75073	hypothetical prote
37	24	92.3	330	2	AD2082	iron(III) dicitrat
38	24	92.3	333	2	AB2130	iron(III) dicitrat
39	24	92.3	337	2	T18708	hypothetical prote
40	24	92.3	346	2	T13837	NADH dehydrogenase
41	24	92.3	352	1	E71092	hypothetical prote
42	24	92.3	356	2	T06756	hypothetical prote
43	24	92.3	386	2	D42528	B23R protein - vac
44	24	92.3	387	2	T26735	hypothetical prote
45	24	92.3	389	2	AH3003	penicillin-binding

ALIGNMENTS

RESULT 1
A47154
conserved hypothetical protein ylxM - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C;Accession: A47154; A69882
R;Honda, K.; Nakamura, K.; Nishiguchi, M.; Yamane, K.
J. Bacteriol. 175, 4885-4894, 1993
A;Title: Cloning and characterization of a Bacillus subtilis gene encoding a homolog
A;Reference number: A47154; MUID:93328695
A;Accession: A47154
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-110 <HON>
A;Cross-references: GB:D14356; NID:g439700; PIDN:BAA22221.1; PID:g2424968
A;Note: sequence extracted from NCBI backbone (NCBIN:135652, NCBIP:135653)
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033
A;Accession: A69882
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-110 <KUN>
A;Cross-references: GB:Z99112; GB:AL009126; NID:g2633902; PIDN:CAB13470.1; PID:g26339
A;Experimental source: strain 168
C;Genetics:
A;Gene: ylxM
C;Superfamily: hypothetical protein A05_orf102

Query Match 100.0%; Score 26; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 KLLLLK 6
           ||||||
Db      69 KLLLLK 74
```

RESULT 2
F81868
probable proline-rich repeat protein NMA1723 [imported] - Neisseria meningitidis (strain
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: F81868
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556
A;Accession: F81868
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <PAR>
A;Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB84951.1; PID:g738036
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1723

Query Match 100.0%; Score 26; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 50 KLLLLK 55

RESULT 3
B72621
hypothetical protein APE1427 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: B72621
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339
A;Accession: B72621
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-126 <KAW>
A;Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAA80424.1; PID:g5105110
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1427
C;Superfamily: Aeropyrum pernix hypothetical protein APE1427

Query Match 100.0%; Score 26; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 62 KLLLLK 67

RESULT 4
C87309
peptidyl-tRNA hydrolase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 14-Sep-2001
C;Accession: C87309
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87309
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-205 <STO>
A;Cross-references: GB:AE005673; NID:g13421661; PIDN:AAK22471.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC0484
C;Superfamily: peptidyl-tRNA hydrolase

Query Match 100.0%; Score 26; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 59 KLLLLK 64

RESULT 5
AB2461
hypothetical protein alr5242 [imported] - Anabaena sp. (strain PCC 7120)
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AB2461
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2461
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-277 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB76941.1; PID:g17134381; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr5242

Query Match 100.0%; Score 26; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 66 KLLLLK 71

RESULT 6
A70212
conserved hypothetical protein BBA41 - Lyme disease spirochete plasmid A/lp54
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C;Accession: A70212
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943
A;Accession: A70212
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-305 <KLE>
A;Cross-references: GB:AE000790; NID:g2690224; PIDN:AAC66251.1; PID:g2690250; TIGR:BB
A;Experimental source: strain B31
C;Genetics:
A;Genome: plasmid

```
  Query Match           100.0%;  Score 26;  DB 2;  Length 305;
  Best Local Similarity  100.0%;  Pred. No. 94;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    257 KLLLLK 262


RESULT   7
S51395
hypothetical protein YLR257w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L8479.9
C;Species: Saccharomyces cerevisiae
C;Date: 05-May-1995 #sequence_revision 12-May-1995 #text_change 05-Nov-1999
C;Accession: S51395
R;Miller, N.
submitted to the EMBL Data Library, November 1994
A;Description: The sequence of S. cerevisiae cosmid 8479.
A;Reference number: S51395
A;Accession: S51395
A;Molecule type: DNA
A;Residues: 1-321 <MIL>
A;Cross-references: EMBL:U17244; NID:g577171; PIDN:AAB67379.1; PID:g577180; GSPDB:GN0001
C;Genetics:
A;Gene: MIPS:YLR257w
A;Map position: 12R


  Query Match           100.0%;  Score 26;  DB 2;  Length 321;
  Best Local Similarity  100.0%;  Pred. No. 99;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    109 KLLLLK 114


RESULT   8
A81745
conserved hypothetical protein TC0068 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
C;Accession: A81745
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255
A;Accession: A81745
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-392 <TET>
A;Cross-references: GB:AE002275; GB:AE002160; NID:g7190108; PIDN:AAF38951.1; PID:g719010
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0068
C;Superfamily: conserved hypothetical protein CP0072


  Query Match           100.0%;  Score 26;  DB 2;  Length 392;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    113 KLLLLK 118


RESULT   9
B71482
hypothetical protein CT696 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 18-Aug-2000
C;Accession: B71482
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A;Reference number: A71570; MUID:99000809
A;Accession: B71482
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <ARN>
A;Cross-references: GB:AE001340; GB:AE001273; NID:g3329147; PIDN:AAC68291.1; PID:g332
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: CT696
C;Superfamily: conserved hypothetical protein CP0072


  Query Match           100.0%;  Score 26;  DB 2;  Length 403;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    124 KLLLLK 129


RESULT  10
B81910
probable ABC-transporter ATP-binding protein NMA1409 [imported] - Neisseria meningiti
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: B81910
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491
A;Reference number: A81775; MUID:20222556
A;Accession: B81910
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-542 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84649.1; PID:g738
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1409
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog


  Query Match           100.0%;  Score 26;  DB 2;  Length 542;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    462 KLLLLK 467


RESULT  11
E81105
ABC transporter, ATP-binding protein NMB1240 [imported] - Neisseria meningitidis (str
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: E81105
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755
A;Accession: E81105
A;Status: preliminary
A;Molecule type: DNA
```

A;Residues: 1-542 <TET>
A;Cross-references: GB:AE002472; GB:AE002098; NID:g7226475; PIDN:AAF41621.1; PID:g722648
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1240
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 100.0%; Score 26; DB 2; Length 542;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 462 KLLLLK 467

RESULT 12
T18437
hypothetical protein C0405c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C;Accession: T18437
R;Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z18935
A;Accession: T18437
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-608 <LAW>
A;Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325391; PIDN:CAB11116.1
C;Genetics:
A;Map position: 3
A;Note: C0405c

Query Match 100.0%; Score 26; DB 2; Length 608;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 500 KLLLLK 505

RESULT 13
A39961
choline O-acetyltransferase (EC 2.3.1.6) precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 05-May-2000
C;Accession: A39961; PH1572; A28047; I46574
R;Berrard, S.; Brice, A.; Lottspeich, F.; Braun, A.; Barde, Y.A.; Mallet, J.
Proc. Natl. Acad. Sci. U.S.A. 84, 9280-9284, 1987
A;Title: cDNA cloning and complete sequence of porcine choline acetyltransferase: in vit
A;Reference number: A39961; MUID:88097472
A;Accession: A39961
A;Molecule type: mRNA
A;Residues: 1-641 <BER>
A;Cross-references: GB:J03021; NID:g164377; PIDN:AAA31000.1; PID:g164378
R;Hersh, L.B.; Kong, C.F.; Sampson, C.; Mues, G.; Li, Y.P.; Fisher, A.; Hilt, D.; Baetge
J. Neurochem. 61, 306-314, 1993
A;Title: Comparison of the promoter region of the human and porcine choline acetyltransf
A;Reference number: PH1571; MUID:93294599
A;Accession: PH1572
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-22 <HER>
R;Braun, A.; Barde, Y.A.; Lottspeich, F.; Mewes, W.; Thoenen, H.
J. Neurochem. 48, 16-21, 1987
A;Title: N-terminal sequence of pig brain choline acetyltransferase purified by a rapid
A;Reference number: A28047; MUID:87085562
A;Accession: A28047
A;Molecule type: protein
A;Residues: 2-12 <BRA>
R;Berrard, S.; Brice, A.E.; Mallet, J.
Brain Res. Bull. 22, 147-153, 1989
A;Title: Molecular genetic approach to the study of mammalian choline acetyltransfera
A;Reference number: I46574; MUID:89229974
A;Accession: I46574
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-641 <BE2>
A;Cross-references: GB:M27736; NID:g164414; PIDN:AAA31015.1; PID:g164415
C;Comment: This enzyme is responsible for the biosynthesis of the neurotransmitter ac
C;Superfamily: carnitine O-acetyltransferase
C;Keywords: acyltransferase; coenzyme A

Query Match 100.0%; Score 26; DB 2; Length 641;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 489 KLLLLK 494

RESULT 14
S32961
hypothetical protein YBR259w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YBR1727
C;Species: Saccharomyces cerevisiae
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 04-Mar-2000
C;Accession: S32961; S46140
R;Doignon, F.; Biteau, N.; Crouzet, M.; Aigle, M.
Yeast 9, 189-199, 1993
A;Title: The complete sequence of a 19,482 bp segment located on the right arm of chr
A;Reference number: S29348; MUID:93220397
A;Accession: S32961
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-688 <DOI>
A;Cross-references: EMBL:X70529; NID:g1907246; PIDN:CAA49923.1; PID:g296558
R;Aigle, M.; Baclet, M.C.; Barthe, C.; Biteau, N.; Crouzet, M.; Doignon, F.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45940
A;Accession: S46140
A;Molecule type: DNA
A;Residues: 1-688 <AIG>
A;Cross-references: EMBL:Z36128; NID:g536684; PIDN:CAA85222.1; PID:g536685; MIPS:YBR2
C;Genetics:
A;Map position: 2R
C;Superfamily: Saccharomyces cerevisiae hypothetical protein YBR259w

Query Match 100.0%; Score 26; DB 2; Length 688;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 655 KLLLLK 660

RESULT 15
A60202
choline O-acetyltransferase (EC 2.3.1.6) - human
N;Alternate names: choline acetylase
C;Species: Homo sapiens (man)
C;Date: 10-Nov-1992 #sequence_revision 13-Mar-1997 #text_change 03-Nov-2000
C;Accession: I52631; A60202; S24416; S14483; PH1571; T01786; PC4342; PC4344; PC4343
R;Oda, Y.; Nakanishi, I.; Deguchi, T.
Brain Res. Mol. Brain Res. 16, 287-294, 1992
A;Title: A complementary DNA for human choline acetyltransferase induces two forms of
A;Reference number: I52631; MUID:93180642
A;Accession: I52631

A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-748 <RES>
A;Cross-references: GB:S56138; NID:g301095; PID:g301096
R;Hersh, L.B.; Takane, K.; Gylys, K.; Moomaw, C.; Slaughter, C.
J. Neurochem. 51, 1843-1845, 1988
A;Title: Conservation of amino acid sequences between human and porcine choline acetyltr
A;Reference number: A60202; MUID:89036242
A;Accession: A60202
A;Molecule type: protein
A;Residues: 'XX',163-182;271-295;340-352;376-382;404-415;550-559;572-583;620-632;644-648
R;Toussaint, J.L.; Geoffroy, V.; Schmitt, M.; Werner, A.; Garnier, J.M.; Simoni, P.; Ken
Genomics 12, 412-416, 1992
A;Title: Human choline acetyltransferase (CHAT): partial gene sequence and potential con
A;Reference number: S24416; MUID:92155737
A;Accession: S24416
A;Molecule type: DNA
A;Residues: 109-150,'Q',152-232 <TOU>
A;Cross-references: EMBL:X56585; NID:g29938; PIDN:CAA39923.1; PID:g29939
R;Cervini, R.; Rocchi, M.; DiDonato, S.; Finocchiaro, G.
submitted to the EMBL Data Library, January 1991
A;Description: Isolation and sub-chromosomal localization of a DNA fragment of the human
A;Reference number: S14483
A;Accession: S14483
A;Molecule type: DNA
A;Residues: 688-738 <CER>
A;Cross-references: EMBL:X56879; NID:g29940; PID:g29941
R;Hersh, L.B.; Kong, C.F.; Sampson, C.; Mues, G.; Li, Y.P.; Fisher, A.; Hilt, D.; Baetge
J. Neurochem. 61, 306-314, 1993
A;Title: Comparison of the promoter region of the human and porcine choline acetyltransf
A;Reference number: PH1571; MUID:93294599
A;Accession: PH1571
A;Status: translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-95 <HE2>
A;Experimental source: GB:L33837
A;Note: GenBank sequence g505335 (accession L33837) is missing one nucleotide C at posit
R;Lorenzi, M.V.; Trinidad, A.C.; Zhang, R.; Strauss, W.L.
DNA Cell Biol. 11, 593-603, 1992
A;Title: Two mRNAs are transcribed from the human gene for choline acetyltransferase.
A;Reference number: Z14429; MUID:93000480
A;Accession: T01786
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 111-119,'T',121-260,'GQ',263-391,'A',393-395,'L',397-433,'G',435-528,'S',530
A;Cross-references: EMBL:S45018; NID:g257109; PIDN:AAB23557.2; PID:g5705927
R;Misawa, H.; Matsuura, J.; Oda, Y.; Takahashi, R.; Deguchi, T.
Mol. Brain Res. 44, 323-333, 1997
A;Title: Human choline acetyltransferase mRNAs with different 5'-region produce a 69-kDa
A;Reference number: PC4342; MUID:97225904
A;Accession: PC4342
A;Molecule type: mRNA
A;Residues: 119-167,'E',169-256 <MIS>
A;Cross-references: DDBJ:D82340
A;Experimental source: brain
A;Accession: PC4344
A;Molecule type: mRNA
A;Residues: 119-152 <MI3>
A;Cross-references: DDBJ:D82341; NID:g1906787; PIDN:BAA18945.1; PID:d1019686; PID:g19067
A;Experimental source: brain
A;Accession: PC4343
A;Molecule type: mRNA
A;Residues: 119-152 <MI2>
A;Cross-references: DDBJ:D82342; NID:g1906789; PIDN:BAA18946.1; PID:d1019687; PID:g19067
A;Experimental source: brain
C;Comment: This enzyme is responsible for the biosynthesis of the neurotransmitter acety
C;Comment: This enzyme is involved in the synthesis of the neurotransmitter acetylcholin
C;Comment: This enzyme is involved in the synthesis of the neurotransmitter acetylcholin
C;Genetics:
A;Gene: GDB:CHAT
A;Cross-references: GDB:119775; OMIM:118490
A;Map position: 10q11.2-10q11.2
A;Introns: 129/3; 193/3

C;Superfamily: carnitine O-acetyltransferase
C;Keywords: acyltransferase; coenzyme A

Query Match 100.0%; Score 26; DB 2; Length 748;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLLK 6
||||||
Db 596 KLLLLK 601

Search completed: June 17, 2002, 12:42:59
Job time: 254 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:44:45 ; Search time 21.35 Seconds
(without alignments)
10.881 Million cell updates/sec

Title: US-09-367-714A-28
Perfect score: 26
Sequence: 1 KLLLLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	100.0	110	1	YLXM_BACSU	P37104 bacillus su
2	26	100.0	194	1	YZ10_AQUAE	O66404 aquifex aeo
3	26	100.0	392	1	Y068_CHLMU	Q9pln1 chlamydia m
4	26	100.0	392	1	Y696_CHLTR	O84702 chlamydia t
5	26	100.0	640	1	CLAT_PIG	P13222 sus scrofa
6	26	100.0	688	1	YB9F_YEAST	P38338 saccharomyc
7	26	100.0	748	1	CLAT_HUMAN	P28329 homo sapien
8	26	100.0	2431	1	POLN_SFV	P08411 semliki for
9	24	92.3	80	1	Y385_BUCAI	P57465 buchnera ap
10	24	92.3	93	1	YSC1_THEFL	P25124 thermus aqu
11	24	92.3	165	1	Y245_MYCGE	P47487 mycoplasma
12	24	92.3	239	1	UL24_HSV7J	P52386 human herpe
13	24	92.3	260	1	HPA1_HELPY	Q48264 helicobacte
14	24	92.3	386	1	VC17_VACCC	P21101 vaccinia vi
15	24	92.3	391	1	DNM2_HUMAN	O14717 homo sapien
16	24	92.3	399	1	XYLR_CALSR	P40981 caldicellul
17	24	92.3	401	1	ISPE_LYCES	P93841 lycopersico
18	24	92.3	407	1	MNDA_HUMAN	P41218 homo sapien
19	24	92.3	439	1	AKR_ARATH	Q05753 arabidopsis
20	24	92.3	469	1	C39A_HUMAN	Q9ny15 homo sapien
21	24	92.3	565	1	HXB1_HAEIN	P44601 haemophilus
22	24	92.3	565	1	HXB2_HAEIN	P45356 haemophilus
23	24	92.3	1005	1	YCF1_OENBE	P31563 oenothera b
24	24	92.3	1066	1	KL61_DROME	P46863 drosophila
25	24	92.3	1139	1	KPC1_TRIRE	Q99014 trichoderma
26	24	92.3	1804	1	YFA7_YEAST	P43583 saccharomyc
27	24	92.3	2052	1	FYV1_MOUSE	Q9z1t6 mus musculu
28	23	88.5	25	1	SPIG_PSEUS	P82357 pseudacanth
29	23	88.5	88	1	Y007_STRPN	Q97tc9 streptococc
30	23	88.5	108	1	RK21_CYACA	O19884 cyanidium c
31	23	88.5	121	1	RL3_ACTAC	P94799 actinobacil
32	23	88.5	123	1	YB8F_YEAST	P38357 saccharomyc
33	23	88.5	130	1	CW36_YEAST	P25603 saccharomyc
34	23	88.5	156	1	SOXR_PSEAE	Q51506 pseudomonas
35	23	88.5	178	1	CASK_RAT	P04468 rattus norv
36	23	88.5	190	1	YEIP_ECOLI	P33028 escherichia
37	23	88.5	208	1	RL3_HAEIN	P44344 haemophilus
38	23	88.5	265	1	HEM4_HUMAN	P10746 homo sapien
39	23	88.5	265	1	HEM4_MOUSE	P51163 mus musculu
40	23	88.5	362	1	MTM1_MICAM	P50190 microbacter
41	23	88.5	388	1	ARP2_ACACA	P53487 acanthamoeb
42	23	88.5	394	1	ARP2_CHICK	P53488 gallus gall
43	23	88.5	394	1	ARP2_HUMAN	O15142 homo sapien
44	23	88.5	395	1	ARP2_DROME	P45888 drosophila
45	23	88.5	399	1	XYLR_ANATH	Q44406 anaerocellu

ALIGNMENTS

RESULT 1
YLXM_BACSU
ID YLXM_BACSU STANDARD; PRT; 110 AA.
AC P37104;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 13.2 kDa protein in ftsY-ffh intergenic region.
GN YLXM.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=93328695; PubMed=8335643;
RA Honda K., Nakamura K., Nishiguchi M., Yamane K.;
RT "Cloning and characterization of a Bacillus subtilis gene encoding a
RT homolog of the 54-kilodalton subunit of mammalian signal recognition
RT particle and Escherichia coli Ffh.";
RL J. Bacteriol. 175:4885-4894(1993).
CC -!- FUNCTION: NOT KNOWN. COULD TAKE PART IN THE SIGNAL RECOGNITION
CC PARTICLE (SRP) PATHWAY. THIS IS INFERRED DUE TO THE CONSERVATION
CC OF ITS GENETIC PROXIMITY TO FTSY/FFH. MAY BE A REGULATORY PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE UPF0122 FAMILY.
CC --

CC --
DR EMBL; D14356; BAA22221.1; -.
DR EMBL; Z99112; CAB13470.1; -.
DR PIR; A47154; A47154.
DR SubtiList; BG10829; ylxM.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 110 AA; 13165 MW; 8DFFEE0A940CD6B9 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 KLLLLK 6
           ||||||
Db      69 KLLLLK 74
```

RESULT 2
YZ10_AQUAE
ID YZ10_AQUAE STANDARD; PRT; 194 AA.
AC O66404;

```
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein AA10.
GN   AA10.
OS   Aquifex aeolicus.
OG   Plasmid ecel.
OC   Bacteria; Aquificales; Aquificaceae; Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
CC   -!- SIMILARITY: TO A.AEOLICUS AQ_423.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000667; AAC07956.1; -.
KW   Hypothetical protein; Plasmid; Complete proteome.
SQ   SEQUENCE   194 AA;  23393 MW;  816FF102AD5E317D CRC64;

  Query Match           100.0%;  Score 26;  DB 1;  Length 194;
  Best Local Similarity  100.0%;  Pred. No. 23;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db     60 KLLLLK 65


RESULT   3
Y068_CHLMU
ID   Y068_CHLMU     STANDARD;      PRT;   392 AA.
AC   Q9PLN1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein TC0068.
GN   TC0068.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0675/CT696/TC0068
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002274; AAF38951.1; -.
DR   TIGR; TC0068; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   392 AA;  45694 MW;  8C7257499347A7C4 CRC64;

  Query Match           100.0%;  Score 26;  DB 1;  Length 392;
  Best Local Similarity  100.0%;  Pred. No. 45;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    113 KLLLLK 118


RESULT   4
Y696_CHLTR
ID   Y696_CHLTR     STANDARD;      PRT;   392 AA.
AC   O84702;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein CT696.
GN   CT696.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/CX;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0675/CT696/TC0068
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001340; AAC68291.2; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   392 AA;  45712 MW;  4EC06DEE24CF8E69 CRC64;

  Query Match           100.0%;  Score 26;  DB 1;  Length 392;
  Best Local Similarity  100.0%;  Pred. No. 45;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    113 KLLLLK 118


RESULT   5
CLAT_PIG
ID   CLAT_PIG       STANDARD;      PRT;   640 AA.
AC   P13222;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
```

```
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Choline O-acetyltransferase (EC 2.3.1.6) (CHOACTase) (Choline
DE   acetylase) (ChAT).
GN   CHAT.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ventral spinal cord;
RX   MEDLINE=88097472; PubMed=3480542;
RA   Berrard S., Brice A., Lottspeich F., Braun A., Barde Y.-A., Mallet J.;
RT   "cDNA cloning and complete sequence of porcine choline
RT   acetyltransferase: in vitro translation of the corresponding RNA
RT   yields an active protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:9280-9284(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ventral spinal cord;
RX   MEDLINE=89229974; PubMed=2713713;
RA   Berrard S., Brice A., Mallet J.;
RT   "Molecular genetic approach to the study of mammalian choline
RT   acetyltransferase.";
RL   Brain Res. Bull. 22:147-153(1989).
RN   [3]
RP   SEQUENCE OF 1-11.
RC   TISSUE=Brain;
RX   MEDLINE=87085562; PubMed=3794697;
RA   Braun A., Barde Y.-A., Lottspeich F., Mewes H.-W., Thoenen H.;
RT   "N-terminal sequence of pig brain choline acetyltransferase purified
RT   by a rapid procedure.";
RL   J. Neurochem. 48:16-21(1987).
CC   -!- FUNCTION: Catalyzes the reversible synthesis of acetylcholine
CC       (ACh) from acetyl CoA and choline at cholinergic synapses.
CC   -!- CATALYTIC ACTIVITY: Acetyl-CoA + choline = CoA + O-acetylcholine.
CC   -!- SIMILARITY: BELONGS TO THE CARNITINE/CHOLINE ACETYLTRANSFERASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03021; AAA31000.1; -.
DR   EMBL; M27736; AAA31015.1; -.
DR   PIR; A28047; A28047.
DR   PIR; A39961; A39961.
DR   InterPro; IPR000542; Carn_acyltransf.
DR   Pfam; PF00755; Carn_acyltransf; 1.
DR   PROSITE; PS00439; ACYLTRANSF_C_1; 1.
DR   PROSITE; PS00440; ACYLTRANSF_C_2; 1.
KW   Transferase; Acyltransferase; Neurotransmitter biosynthesis.
FT   INIT_MET      0      0
FT   ACT_SITE    333    333       POTENTIAL.
SQ   SEQUENCE   640 AA;  71599 MW;  5ECC27BE8B7CC317 CRC64;


  Query Match             100.0%;  Score 26;  DB 1;  Length 640;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db     488 KLLLLK 493


RESULT   6
YB9F_YEAST
ID   YB9F_YEAST     STANDARD;      PRT;   688 AA.
AC   P38338;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Hypothetical 80.4 kDa protein in POP4-SHM1 intergenic region.
GN   YBR259W OR YBR1727.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=93220397; PubMed=8465606;
RA   Doignon F., Biteau N., Crouzet M., Aigle M.;
RT   "The complete sequence of a 19,482 bp segment located on the right
RT   arm of chromosome II from Saccharomyces cerevisiae.";
RL   Yeast 9:189-199(1993).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70529; CAA49923.1; -.
DR   EMBL; Z36128; CAA85222.1; -.
DR   PIR; S32961; S32961.
DR   SGD; S0000463; YBR259W.
KW   Hypothetical protein.
SQ   SEQUENCE   688 AA;  80426 MW;  0BA84837BD7A4B30 CRC64;


  Query Match             100.0%;  Score 26;  DB 1;  Length 688;
  Best Local Similarity   100.0%;  Pred. No. 78;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db     655 KLLLLK 660


RESULT   7
CLAT_HUMAN
ID   CLAT_HUMAN     STANDARD;      PRT;   748 AA.
AC   P28329; Q16488; Q9BQE1; Q9BQ35; Q9BQ23;
DT   01-DEC-1992 (Rel. 24, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Choline O-acetyltransferase (EC 2.3.1.6) (CHOACTase) (Choline
DE   acetylase) (ChAT).
GN   CHAT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spinal cord;
RX   MEDLINE=93180642; PubMed=1337937;
RA   Oda Y., Nakanishi I., Deguchi T.;
RT   "A complementary DNA for human choline acetyltransferase induces two
RT   forms of enzyme with different molecular weights in cultured cells.";
RL   Brain Res. Mol. Brain Res. 16:287-294(1992).
RN   [2]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, VARIANTS FIMG2 P-210; A-211;
RP   T-305; C-420; K-441; G-482; L-498; L-506 AND H-560, AND VARIANTS T-120
RP   AND G-392.
RX   MEDLINE=21117155; PubMed=11172068;
RA   Ohno K., Tsujino A., Brengman J.M., Harper C.M., Bajzer Z., Udd B.,
RA   Beyring R., Robb S., Kirkham F.J., Engel A.G.;
```

```
RT   "Choline acetyltransferase mutations cause myasthenic syndrome
RT   associated with episodic apnea in humans.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:2017-2022(2001).
RN   [3]
RP   SEQUENCE OF 111-669 FROM N.A.
RX   MEDLINE=93000480; PubMed=1388731;
RA   Lorenzi M.V., Trinidad A.C., Zhang R., Strauss W.L.;
RT   "Two mRNAs are transcribed from the human gene for choline
RT   acetyltransferase.";
RL   DNA Cell Biol. 11:593-603(1992).
RN   [4]
RP   SEQUENCE OF 109-232 FROM N.A.
RX   MEDLINE=92155737; PubMed=1339386;
RA   Toussaint J.L., Geoffroy V., Schmitt M., Werner A., Garnier J.M.,
RA   Simoni P., Kempf J.;
RT   "Human choline acetyltransferase (CHAT): partial gene sequence and
RT   potential control regions.";
RL   Genomics 12:412-416(1992).
RN   [5]
RP   SEQUENCE OF 688-738 FROM N.A.
RC   TISSUE=Lymphocytes;
RX   MEDLINE=92149876; PubMed=1784419;
RA   Cervini R., Rocchi M., DiDonato S., Finocchiaro G.;
RT   "Isolation and sub-chromosomal localization of a DNA fragment of the
RT   human choline acetyltransferase gene.";
RL   Neurosci. Lett. 132:191-194(1991).
CC   -!- FUNCTION: Catalyzes the reversible synthesis of acetylcholine
CC       (ACh) from acetyl CoA and choline at cholinergic synapses.
CC   -!- CATALYTIC ACTIVITY: Acetyl-CoA + choline = CoA + O-acetylcholine.
CC   -!- ALTERNATIVE PRODUCTS: 3 isoforms; M/83 kDa (shown here), S/74 kDa
CC       and R/70 kDa; are produced by alternative splicing.
CC   -!- DISEASE: Defects in CHAT are a cause of familial infantile
CC       myasthenic gravis type 2 (FIMG2) (also known as CMS-EA). FIMG2
CC       patients have myasthenic symptoms since birth or early infancy,
CC       negative tests for anti-AChR antibodies, and abrupt episodic
CC       crises with increased weakness, bulbar paralysis, and apnea
CC       precipitated by undue exertion, fever, or excitement.
CC   -!- SIMILARITY: BELONGS TO THE CARNITINE/CHOLINE ACETYLTRANSFERASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S56138; AAA14245.1; -.
DR   EMBL; AF305907; AAK08953.1; -.
DR   EMBL; AF305906; AAK08950.1; -.
DR   EMBL; AF305894; AAK08950.1; JOINED.
DR   EMBL; AF305895; AAK08950.1; JOINED.
DR   EMBL; AF305896; AAK08950.1; JOINED.
DR   EMBL; AF305897; AAK08950.1; JOINED.
DR   EMBL; AF305898; AAK08950.1; JOINED.
DR   EMBL; AF305899; AAK08950.1; JOINED.
DR   EMBL; AF305900; AAK08950.1; JOINED.
DR   EMBL; AF305901; AAK08950.1; JOINED.
DR   EMBL; AF305902; AAK08950.1; JOINED.
DR   EMBL; AF305903; AAK08950.1; JOINED.
DR   EMBL; AF305904; AAK08950.1; JOINED.
DR   EMBL; AF305905; AAK08950.1; JOINED.
DR   EMBL; AF305908; AAK08954.1; -.
DR   EMBL; AF305906; AAK08951.1; -.
DR   EMBL; AF305894; AAK08951.1; JOINED.
DR   EMBL; AF305895; AAK08951.1; JOINED.
DR   EMBL; AF305896; AAK08951.1; JOINED.
DR   EMBL; AF305897; AAK08951.1; JOINED.
DR   EMBL; AF305898; AAK08951.1; JOINED.
DR   EMBL; AF305899; AAK08951.1; JOINED.
DR   EMBL; AF305900; AAK08951.1; JOINED.
DR   EMBL; AF305901; AAK08951.1; JOINED.
DR   EMBL; AF305902; AAK08951.1; JOINED.
DR   EMBL; AF305903; AAK08951.1; JOINED.
DR   EMBL; AF305904; AAK08951.1; JOINED.
DR   EMBL; AF305905; AAK08951.1; JOINED.
DR   EMBL; AF305909; AAK08955.1; -.
DR   EMBL; AF305906; AAK08952.1; -.
DR   EMBL; AF305894; AAK08952.1; JOINED.
DR   EMBL; AF305895; AAK08952.1; JOINED.
DR   EMBL; AF305896; AAK08952.1; JOINED.
DR   EMBL; AF305897; AAK08952.1; JOINED.
DR   EMBL; AF305898; AAK08952.1; JOINED.
DR   EMBL; AF305899; AAK08952.1; JOINED.
DR   EMBL; AF305900; AAK08952.1; JOINED.
DR   EMBL; AF305901; AAK08952.1; JOINED.
DR   EMBL; AF305902; AAK08952.1; JOINED.
DR   EMBL; AF305903; AAK08952.1; JOINED.
DR   EMBL; AF305904; AAK08952.1; JOINED.
DR   EMBL; AF305905; AAK08952.1; JOINED.
DR   EMBL; S45018; AAB23557.2; -.
DR   EMBL; X56585; CAA39923.1; -.
DR   EMBL; X56879; CAA40201.1; -.
DR   PIR; S14483; S14483.
DR   PIR; S24416; S24416.
DR   MIM; 118490; -.
DR   MIM; 254210; -.
DR   InterPro; IPR000542; Carn_acyltransf.
DR   Pfam; PF00755; Carn_acyltransf; 1.
DR   PROSITE; PS00439; ACYLTRANSF_C_1; 1.
DR   PROSITE; PS00440; ACYLTRANSF_C_2; 1.
KW   Transferase; Acyltransferase; Neurotransmitter biosynthesis;
KW   Alternative splicing; Disease mutation; Polymorphism.
FT   ACT_SITE     442     442       POTENTIAL.
FT   VARIANT      120     120       A -> T.
FT                                  /FTId=VAR_011675.
FT   VARIANT      210     210       L -> P (IN FIMG2; IMPAIRED ACTIVITY).
FT                                  /FTId=VAR_011666.
FT   VARIANT      211     211       P -> A (IN FIMG2; IMPAIRED ACTIVITY).
FT                                  /FTId=VAR_011667.
FT   VARIANT      305     305       I -> T (IN FIMG2; IMPAIRED ACTIVITY).
FT                                  /FTId=VAR_011668.
FT   VARIANT      392     392       A -> G.
FT                                  /FTId=VAR_011676.
FT   VARIANT      420     420       R -> C (IN FIMG2; IMPAIRED ACTIVITY).
FT                                  /FTId=VAR_011669.
FT   VARIANT      441     441       E -> K (IN FIMG2; COMPLETELY LACK
FT                                  ACTIVITY).
FT                                  /FTId=VAR_011670.
FT   VARIANT      482     482       R -> G (IN FIMG2; IMPAIRED ACTIVITY).
FT                                  /FTId=VAR_011671.
FT   VARIANT      498     498       S -> L (IN FIMG2; IMPAIRED ACTIVITY).
FT                                  /FTId=VAR_011672.
FT   VARIANT      506     506       V -> L (IN FIMG2; IMPAIRED ACTIVITY).
FT                                  /FTId=VAR_011673.
FT   VARIANT      560     560       R -> H (IN FIMG2; IMPAIRED ACTIVITY).
FT                                  /FTId=VAR_011674.
FT   VARSPLIC       1      95       MGLRTAKKRGLGGGGKWKREEGGGTRGRREVRPACFLQSGG
FT                                  RGDPGDVGGPAGNPGCSPHPRAATRPPPLPAHTPAHTPEWC
FT                                  GAASAEAAEPRRA -> MWPECRDEALSTV (IN
FT                                  ISOFORM S).
FT   VARSPLIC       1     118       MISSING (IN ISOFORM R).
FT   CONFLICT     151     151       R -> Q (IN REF. 4).
FT   CONFLICT     261     262       GQ -> PE (IN REF. 1).
FT   CONFLICT     396     396       V -> L (IN REF. 3).
FT   CONFLICT     434     434       G -> A (IN REF. 1).
FT   CONFLICT     529     529       C -> S (IN REF. 3).
FT   CONFLICT     567     567       V -> L (IN REF. 3).
FT   CONFLICT     629     630       EL -> DV (IN REF. 3).
FT   CONFLICT     664     664       T -> M (IN REF. 3).
SQ   SEQUENCE   748 AA;  82567 MW;  0D537659F876C497 CRC64;

Query Match              100.0%;  Score 26;  DB 1;  Length 748;
Best Local Similarity    100.0%;  Pred. No. 85;
```

```
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db     596 KLLLLK 601


RESULT   8
POLN_SFV
ID   POLN_SFV       STANDARD;      PRT;  2431 AA.
AC   P08411;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nonstructural polyprotein [Contains: Nonstructural proteins NSP1 TO
DE   NSP4].
OS   Semliki forest virus (SFV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC   Alphavirus.
OX   NCBI_TaxID=11033;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86286581; PubMed=3488539;
RA   Takkinen K.;
RT   "Complete nucleotide sequence of the nonstructural protein genes of
RT   Semliki Forest virus.";
RL   Nucleic Acids Res. 14:5667-5682(1986).
CC   -!- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04129; CAA27741.1; -.
DR   PIR; A23592; MNWVSF.
DR   MEROPS; C09.001; -.
DR   InterPro; IPR002589; Alpp.
DR   InterPro; IPR002620; Peptidase_C9.
DR   InterPro; IPR001788; RNA_dep_RNApol2.
DR   InterPro; IPR000606; Viral_helicase1.
DR   Pfam; PF01661; Alpp; 1.
DR   Pfam; PF01707; Peptidase_C9; 1.
DR   Pfam; PF00978; RNA_dep_RNApol2; 1.
DR   Pfam; PF01443; Viral_helicase1; 1.
DR   SMART; SM00506; Alpp; 1.
KW   Polyprotein; Nonstructural protein; RNA-binding; Helicase.
FT   CHAIN         1    537       NONSTRUCTURAL PROTEIN NSP1.
FT   CHAIN       538   1335       NONSTRUCTURAL PROTEIN NSP2.
FT   CHAIN      1336   1817       NONSTRUCTURAL PROTEIN NSP3.
FT   CHAIN      1818   2431       NONSTRUCTURAL PROTEIN NSP4.
SQ   SEQUENCE   2431 AA;  269286 MW;  1F9EBA1022E3EC5F CRC64;


  Query Match             100.0%;  Score 26;  DB 1;  Length 2431;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db    1863 KLLLLK 1868


RESULT   9
Y385_BUCAI
ID   Y385_BUCAI     STANDARD;      PRT;    80 AA.
AC   P57465;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein BU385.
GN   BU385.
OS   Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS   symbiotic bacterium).
OC   Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX   NCBI_TaxID=118099;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOKYO 1998;
RX   MEDLINE=20445173; PubMed=10993077;
RA   Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT   "Genome sequence of the endocellular bacterial symbiont of aphids
RT   Buchnera sp. APS.";
RL   Nature 407:81-86(2000).
CC   -!- SIMILARITY: BELONGS TO THE BOLA / YRBA FAMILY. STRONG, TO E.COLI
CC       YRBA.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AP001119; BAB13088.1; -.
DR   InterPro; IPR002634; BolA.
DR   Pfam; PF01722; BolA; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   80 AA;  9299 MW;  4AFACA590A038131 CRC64;


  Query Match             92.3%;  Score 24;  DB 1;  Length 80;
  Best Local Similarity   83.3%;  Pred. No. 29;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||:|
Db       7 KLLLIK 12


RESULT  10
YSC1_THEFL
ID   YSC1_THEFL     STANDARD;      PRT;    93 AA.
AC   P25124;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein in SCSB 5'region (ORFA) (Fragment).
OS   Thermus aquaticus (subsp. flavus).
OC   Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX   NCBI_TaxID=274;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33923 / AT-62;
RX   MEDLINE=91238680; PubMed=2034208;
RA   Nishiyama M., Horinouchi S., Beppu T.;
RT   "Characterization of an operon encoding succinyl-CoA synthetase and
RT   malate dehydrogenase from Thermus flavus AT-62 and its expression in
RT   Escherichia coli.";
RL   Mol. Gen. Genet. 226:1-9(1991).
CC   -!- SIMILARITY: BELONGS TO THE YCF81 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X54073; CAA38004.1; -.
```

```
DR   PIR; S15948; S15948.
DR   InterPro; IPR002792; DUF90.
DR   Pfam; PF01938; TRAM; 1.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   93 AA;  10011 MW;  CC3B333389B9CCCF CRC64;
```

```
  Query Match             92.3%;  Score 24;  DB 1;  Length 93;
  Best Local Similarity   83.3%;  Pred. No. 34;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           |||:||
Db      32 KLLILK 37
```

```
RESULT  11
Y245_MYCGE
ID   Y245_MYCGE     STANDARD;      PRT;   165 AA.
AC   P47487;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein MG245.
GN   MG245.
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RX   MEDLINE=96026346; PubMed=7569993;
RA   Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA   Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA   Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA   Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA   Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA   Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT   "The minimal gene complement of Mycoplasma genitalium.";
RL   Science 270:397-403(1995).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U39703; AAC71465.1; -.
DR   TIGR; MG245; -.
DR   InterPro; IPR002698; 5-FTHF_cyc-lig.
DR   Pfam; PF01812; 5-FTHF_cyc-lig; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   165 AA;  19355 MW;  AEC4ADEEE55A7020 CRC64;
```

```
  Query Match             92.3%;  Score 24;  DB 1;  Length 165;
  Best Local Similarity   83.3%;  Pred. No. 59;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||:|||
Db       9 KLMLLK 14
```

```
RESULT  12
UL24_HSV7J
ID   UL24_HSV7J     STANDARD;      PRT;   239 AA.
AC   P52386;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Protein U49.
GN   U49.
OS   Human herpesvirus (type 7 / strain JI) (HHV7).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=57278;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nicholas J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL24,
CC       EHV-1 37, EBV BXRF1, HCMV UL76, ILTV ORF3, AND VZV 35.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U43400; AAC54711.1; -.
DR   InterPro; IPR002580; Herpes_UL24.
DR   Pfam; PF01646; Herpes_UL24; 1.
SQ   SEQUENCE   239 AA;  28568 MW;  F4EBC038679BE6B0 CRC64;
```

```
  Query Match             92.3%;  Score 24;  DB 1;  Length 239;
  Best Local Similarity   83.3%;  Pred. No. 85;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           |:||||
Db      24 KLLLLK 29
```

```
RESULT  13
HPA1_HELPY
ID   HPA1_HELPY     STANDARD;      PRT;   260 AA.
AC   Q48264;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Neuraminyllactose-binding hemagglutinin precursor (N-
DE   acetylneuraminyllactose-binding fibrillar hemagglutinin receptor-
DE   binding subunit) (NLBH) (Flagellar sheath adhesin).
GN   HPAA.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=8826;
RX   MEDLINE=93139035; PubMed=7678592;
RA   Evans D.G., Karjalainen T.K., Evans D.J., Graham D.Y., Lee C.-H.;
RT   "Cloning, nucleotide sequence, and expression of a gene encoding an
RT   adhesin subunit protein of Helicobacter pylori.";
RL   J. Bacteriol. 175:674-683(1993).
CC   -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC       anchor (Probable).
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; X61574; CAA43773.1; -.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Flagella; Outer membrane; Lipoprotein; Signal.
FT   SIGNAL        1     27       BY SIMILARITY.
FT   CHAIN        28    260       NEURAMINYLLACTOSE-BINDING HEMAGGLUTININ.
FT   LIPID        28     28       N-ACYL DIGLYCERIDE (PROBABLE).
FT   DOMAIN      134    139       N-ACETYL-NEURAMINYL-ALPHA(2,3)-LACTOSE
FT                                BINDING MOTIF (POTENTIAL).
SQ   SEQUENCE   260 AA;  29166 MW;  22489598065E7B14 CRC64;


  Query Match             92.3%;  Score 24;  DB 1;  Length 260;
  Best Local Similarity   83.3%;  Pred. No. 93;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           |:||||
Db      57 KILLLK 62


RESULT  14
VC17_VACCC
ID   VC17_VACCC     STANDARD;      PRT;   386 AA.
AC   P21101;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Protein C17/B23.
GN   B23R AND C17L.
OS   Vaccinia virus (strain Copenhagen).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10249;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91021027; PubMed=2219722;
RA   Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA   Paoletti E.;
RT   "The complete DNA sequence of vaccinia virus.";
RL   Virology 179:247-266(1990).
RN   [2]
RP   COMPLETE GENOME.
RA   Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA   Paoletti E.;
RT   "Appendix to 'The complete DNA sequence of vaccinia virus'.";
RL   Virology 179:517-563(1990).
CC   -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M35027; AAA48222.1; -.
DR   EMBL; M35027; AAA47979.1; -.
DR   PIR; D42528; D42528.
DR   PIR; F42502; F42502.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 3.
DR   SMART; SM00248; ANK; 1.
DR   PROSITE; PS50088; ANK_REPEAT; FALSE_NEG.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       59     91       ANK 1.
FT   REPEAT       95    126       ANK 2.
FT   REPEAT      210    245       ANK 3.
FT   REPEAT      249    280       ANK 4.
FT   REPEAT      292    321       ANK 5.
FT   REPEAT      349    381       ANK 6.
SQ   SEQUENCE   386 AA;  44941 MW;  D553A134C9317A42 CRC64;


  Query Match             92.3%;  Score 24;  DB 1;  Length 386;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||:|
Db     269 KLLLIK 274


RESULT  15
DNM2_HUMAN
ID   DNM2_HUMAN     STANDARD;      PRT;   391 AA.
AC   O14717;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   DNA (cytosine-5)-methyltransferase-like protein 2 (DNA
DE   methyltransferase homolog HsaIIP) (DNA MTase homolog HsaIIP)
DE   (M.HsaIIP) (PuMet).
GN   DNMT2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98087580; PubMed=9425235;
RA   Yoder J.A., Bestor T.H.;
RT   "A candidate mammalian DNA methyltransferase related to pmt1p of
RT   fission yeast.";
RL   Hum. Mol. Genet. 7:279-284(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98258972; PubMed=9599025;
RA   Van den Wyngaert I., Sprengel J., Kass S.U., Luyten W.H.;
RT   "Cloning and analysis of a novel human putative DNA
RT   methyltransferase.";
RL   FEBS Lett. 426:283-289(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Bird C.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX   MEDLINE=20580737; PubMed=11139614;
RA   Dong A., Yoder J.A., Zhang X., Zhou L., Bestor T.H., Cheng X.;
RT   "Structure of human DNMT2, an enigmatic DNA methyltransferase homolog
RT   that displays denaturant-resistant binding to DNA.";
RL   Nucleic Acids Res. 29:439-448(2001).
CC   -!- FUNCTION: Seems not be active as a DNA methyltransferase. Its
CC       strong binding to DNA suggests that it may mark specific sequences
CC       in the genome by binding to DNA through the specific target-
CC       recognizing motif.
CC   -!- TISSUE SPECIFICITY: Ubiquitous.
CC   -!- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF012128; AAC51939.1; -.
DR   EMBL; AJ223333; CAA11272.1; -.
DR   EMBL; AL133415; CAB87964.1; -.
DR   PDB; 1G55; 17-JAN-01.
DR   REBASE; 3241; M.HsaIIP.
DR   MIM; 602478; -.
DR   InterPro; IPR001525; C5_DNA_meth.
```

```
DR   Pfam; PF00145; DNA_methylase; 2.
DR   PRINTS; PR00105; C5METTRFRASE.
DR   PROSITE; PS00094; C5_MTASE_1; FALSE_NEG.
DR   PROSITE; PS00095; C5_MTASE_2; 1.
KW   DNA-binding; 3D-structure.
FT   ACT_SITE     79      79       BY SIMILARITY.
SQ   SEQUENCE   391 AA;  44596 MW;  BCA549E4EB2E6950 CRC64;


  Query Match             92.3%;  Score 24;  DB 1;  Length 391;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
             |||:||
Db     334 KLLILK 339


Search completed: June 17, 2002, 12:44:46
Job time: 301 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:44:18 ; Search time 73.61 Seconds
(without alignments)
14.101 Million cell updates/sec

Title: US-09-367-714A-28
Perfect score: 26
Sequence: 1 KLLLLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	100.0	58	5	Q95T00	Q95t00 drosophila
2	26	100.0	110	16	Q9JTL6	Q9jtl6 neisseria m
3	26	100.0	126	17	Q9YC24	Q9yc24 aeropyrum p
4	26	100.0	146	3	Q9P7A2	Q9p7a2 schizosacch
5	26	100.0	205	16	Q9AAV9	Q9aav9 caulobacter
6	26	100.0	265	8	Q9MG96	Q9mg96 chrysodidym
7	26	100.0	305	16	O50934	O50934 borrelia bu
8	26	100.0	321	3	Q06146	Q06146 saccharomyc
9	26	100.0	482	8	Q9G8N5	Q9g8n5 naegleria g
10	26	100.0	542	16	Q9JZ89	Q9jz89 neisseria m
11	26	100.0	542	16	Q9JUB3	Q9jub3 neisseria m
12	26	100.0	544	3	Q9HGZ7	Q9hgz7 schizosacch
13	26	100.0	547	2	Q93N56	Q93n56 coxiella bu
14	26	100.0	560	5	Q9TWB0	Q9twb0 plasmodium
15	26	100.0	576	5	Q95TP6	Q95tp6 drosophila
16	26	100.0	608	5	O77331	O77331 plasmodium
17	26	100.0	614	12	O41268	O41268 semliki for
18	26	100.0	684	12	Q91MW5	Q91mw5 lumpy skin
19	26	100.0	841	5	Q9U6X3	Q9u6x3 drosophila
20	26	100.0	1049	12	Q83611	Q83611 ectromelia
21	26	100.0	1096	5	Q9NGX1	Q9ngx1 entamoeba h
22	26	100.0	1173	11	Q99MZ9	Q99mz9 mus musculu
23	26	100.0	1441	10	O24367	O24367 spirodela p
24	26	100.0	1855	5	Q9TX75	Q9tx75 plasmodium
25	26	100.0	1855	5	Q9BHN0	Q9bhn0 plasmodium
26	26	100.0	1912	5	Q9U0H1	Q9u0h1 plasmodium
27	26	100.0	2432	12	Q9QBM1	Q9qbm1 semliki for
28	24	92.3	41	8	O20189	O20189 chlorella v
29	24	92.3	80	10	Q9LGL0	Q9lgl0 oryza sativ
30	24	92.3	90	8	Q35715	Q35715 romanomermi
31	24	92.3	95	16	Q930X5	Q930x5 rhizobium m
32	24	92.3	128	10	Q9XEK3	Q9xek3 dendrobium
33	24	92.3	146	16	Q9JR77	Q9jr77 neisseria m
34	24	92.3	167	2	Q9RNC2	Q9rnc2 helicobacte
35	24	92.3	187	4	Q96SL4	Q96sl4 homo sapien
36	24	92.3	201	16	Q92AN1	Q92an1 listeria in
37	24	92.3	215	16	O84255	O84255 chlamydia t
38	24	92.3	215	16	Q9PKE1	Q9pke1 chlamydia m
39	24	92.3	217	2	Q9RNC4	Q9rnc4 helicobacte
40	24	92.3	217	2	Q9RNC3	Q9rnc3 helicobacte
41	24	92.3	217	16	O25359	O25359 helicobacte
42	24	92.3	217	16	Q9ZLJ3	Q9zlj3 helicobacte
43	24	92.3	233	8	Q9MTD5	Q9mtd5 toxoplasma
44	24	92.3	236	16	Q98EW8	Q98ew8 rhizobium l
45	24	92.3	239	16	Q9PMD0	Q9pmd0 campylobact

ALIGNMENTS

RESULT 1
Q95T00
ID Q95T00 PRELIMINARY; PRT; 58 AA.
AC Q95T00;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE LD32080P.
GN CHARYBDE.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY060401; AAL25440.1; -.
SQ SEQUENCE 58 AA; 7029 MW; 4CE54127CB70AA98 CRC64;

Query Match 100.0%; Score 26; DB 5; Length 58;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 KLLLLK 6
          ||||||
Db     53 KLLLLK 58
```

RESULT 2
Q9JTL6
ID Q9JTL6 PRELIMINARY; PRT; 110 AA.

```
AC   Q9JTL6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE PROLINE-RICH REPEAT PROTEIN.
GN   NMA1723.
OS   Neisseria meningitidis (serogroup A).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
DR   EMBL; AL162756; CAB84951.1; -.
KW   Complete proteome.
SQ   SEQUENCE   110 AA;  12714 MW;  F8ED83151FC34CB8 CRC64;

  Query Match             100.0%;  Score 26;  DB 16;  Length 110;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db     50 KLLLLK 55


RESULT   3
Q9YC24
ID   Q9YC24      PRELIMINARY;      PRT;   126 AA.
AC   Q9YC24;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 14.2 KDA PROTEIN APE1427.
GN   APE1427.
OS   Aeropyrum pernix.
OC   Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC   Aeropyrum.
OX   NCBI_TaxID=56636;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1;
RX   MEDLINE=99310339; PubMed=10382966;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA   Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA   Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA   Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA   Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA   Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT   "Complete genome sequence of an aerobic hyper-thermophilic
RT   crenarchaeon, Aeropyrum pernix K1.";
RL   DNA Res. 6:83-101(1999).
DR   EMBL; AP000061; BAA80424.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   126 AA;  14192 MW;  B766FEF18B135029 CRC64;

  Query Match             100.0%;  Score 26;  DB 17;  Length 126;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db     62 KLLLLK 67


RESULT   4
Q9P7A2
ID   Q9P7A2      PRELIMINARY;      PRT;   146 AA.
AC   Q9P7A2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CYTOPLASMIC DYNEIN INTERMEDIATE CHAIN (FRAGMENT).
GN   SPBP35G2.01C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   Seeger K., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   McCombie W.R.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL163702; CAB87363.1; -.
FT   NON_TER     146    146
SQ   SEQUENCE   146 AA;  16333 MW;  5D283F44B9426B13 CRC64;

  Query Match             100.0%;  Score 26;  DB 3;  Length 146;
  Best Local Similarity   100.0%;  Pred. No. 78;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db     15 KLLLLK 20


RESULT   5
Q9AAV9
ID   Q9AAV9      PRELIMINARY;      PRT;   205 AA.
AC   Q9AAV9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PEPTIDYL-TRNA HYDROLASE.
GN   CC0484.
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC   Caulobacter.
OX   NCBI_TaxID=69394;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA   Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA   Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR   EMBL; AE005721; AAK22471.1; -.
DR   HSSP; P23932; 2PTH.
DR   TIGR; CC0484; -.
DR   InterPro; IPR001328; Pept_tRNA_hydro.
DR   Pfam; PF01195; Pept_tRNA_hydro; 1.
DR   ProDom; PD005324; Pept_tRNA_hydro; 1.
```

```
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   205 AA;  22472 MW;  F64F92A527D6B8F6 CRC64;


  Query Match           100.0%;  Score 26;  DB 16;  Length 205;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db     59 KLLLLK 64


RESULT   6
Q9MG96
ID   Q9MG96      PRELIMINARY;      PRT;   265 AA.
AC   Q9MG96;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RIBOSOMAL PROTEIN S3.
GN   RPS3.
OS   Chrysodidymus synuroideus.
OG   Mitochondrion.
OC   Eukaryota; stramenopiles; Chrysophyceae; Synurales; Chrysodidymus.
OX   NCBI_TaxID=47573;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20330374; PubMed=10871400;
RA   Chesnick J.M., Goff M., Graham J., Ocampo C., Lang B.F., Seif E.,
RA   Burger G.;
RT   "The mitochondrial genome of the stramenopile alga Chrysodidymus
RT   synuroideus. Complete sequence, gene content and genome
RT   organization.";
RL   Nucleic Acids Res. 28:2512-2518(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Burger G.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF222718; AAF36953.1; -.
DR   InterPro; IPR001351; Ribosomal_S3.
DR   Pfam; PF00189; Ribosomal_S3_C; 1.
KW   Mitochondrion.
SQ   SEQUENCE   265 AA;  31348 MW;  D985AB230A78220E CRC64;


  Query Match           100.0%;  Score 26;  DB 8;  Length 265;
  Best Local Similarity  100.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db     99 KLLLLK 104


RESULT   7
O50934
ID   O50934      PRELIMINARY;      PRT;   305 AA.
AC   O50934;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   CONSERVED HYPOTHETICAL PROTEIN.
GN   BBA41.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp54.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
DR   EMBL; AE000790; AAC66251.1; -.
DR   TIGR; BBA41; -.
KW   Plasmid; Complete proteome.
SQ   SEQUENCE   305 AA;  34953 MW;  8A57A4FDF9972F9F CRC64;


  Query Match           100.0%;  Score 26;  DB 16;  Length 305;
  Best Local Similarity  100.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    257 KLLLLK 262


RESULT   8
Q06146
ID   Q06146      PRELIMINARY;      PRT;   321 AA.
AC   Q06146;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CHROMOSOME XII COSMID 8479.
GN   YLR257W OR L8479.9.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C (AB972);
RX   MEDLINE=97313267; PubMed=9169871;
RA   Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
RA   Benes V., Bruckner M., Delius H., Dubois E., Dusterhoft A.,
RA   Entian K.D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA   Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Kotter P.,
RA   Louis E.J., Messenguy F., Mewes H.W., Miosga T., Mostl D.,
RA   Muller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA   Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA   Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA   Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA   Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA   Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL   Nature 387:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C (AB972);
RA   Miller N.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C (AB972);
RA   Waterston R.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C (AB972);
RA   Cherry J.M.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U17244; AAB67379.1; -.
DR   SGD; S0004247; YLR257W.
SQ   SEQUENCE   321 AA;  35998 MW;  E8E1FC17FFB27418 CRC64;
```

```
  Query Match           100.0%;  Score 26;  DB 3;  Length 321;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    109 KLLLLK 114


RESULT   9
Q9G8N5
ID   Q9G8N5       PRELIMINARY;       PRT;   482 AA.
AC   Q9G8N5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   RIBOSOMAL PROTEIN S4.
GN   RPS4.
OS   Naegleria gruberi.
OG   Mitochondrion.
OC   Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OX   NCBI_TaxID=5762;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Burger G., Lang B.F., Nerad T.A., Gray M.W.;
RT   "The mitochondrial genome of the supposedly primitive protist,
RT   Naegleria gruberi.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288092; AAG17819.1; -.
DR   InterPro; IPR002942; S4.
DR   Pfam; PF01479; S4; 1.
DR   SMART; SM00363; S4; 1.
KW   Mitochondrion.
SQ   SEQUENCE   482 AA;  59522 MW;  0A7A5A9E1AB58EAC CRC64;


  Query Match           100.0%;  Score 26;  DB 8;  Length 482;
  Best Local Similarity  100.0%;  Pred. No. 2.3e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db     39 KLLLLK 44


RESULT  10
Q9JZ89
ID   Q9JZ89       PRELIMINARY;       PRT;   542 AA.
AC   Q9JZ89;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ABC TRANSPORTER, ATP-BINDING PROTEIN.
GN   NMB1240.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / SEROGROUP B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002472; AAF41621.1; -.
DR   TIGR; NMB1240; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 2.
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   542 AA;  60778 MW;  1EACB1DC50077CE9 CRC64;


  Query Match           100.0%;  Score 26;  DB 16;  Length 542;
  Best Local Similarity  100.0%;  Pred. No. 2.6e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    462 KLLLLK 467


RESULT  11
Q9JUB3
ID   Q9JUB3       PRELIMINARY;       PRT;   542 AA.
AC   Q9JUB3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ABC-TRANSPORTER ATP-BINDING PROTEIN.
GN   NMA1409.
OS   Neisseria meningitidis (serogroup A).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
DR   EMBL; AL162755; CAB84649.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 2.
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   542 AA;  60723 MW;  715362DFF1AB7527 CRC64;


  Query Match           100.0%;  Score 26;  DB 16;  Length 542;
  Best Local Similarity  100.0%;  Pred. No. 2.6e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    462 KLLLLK 467


RESULT  12
Q9HGZ7
ID   Q9HGZ7       PRELIMINARY;       PRT;   544 AA.
AC   Q9HGZ7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYTOPLASMIC DYNEIN INTERMEDIATE CHAIN.
```

```
GN   SPBC855.01C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   CSHL Advanced;
RT   "Sequence of a 44.9 kb cosmid insert determined during a two week
RT   course.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   McCombie W.R., Lyne M.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR   EMBL; AL391016; CAC01482.1; -.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF00400; WD40; 5.
DR   SMART; SM00320; WD40; 5.
DR   PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR   PROSITE; PS50082; WD_REPEATS_2; 1.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   544 AA;  60960 MW;  1057A3C5435F3E6D CRC64;


  Query Match            100.0%;  Score 26;  DB 3;  Length 544;
  Best Local Similarity  100.0%;  Pred. No. 2.6e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db     15 KLLLLK 20


RESULT  13
Q93N56
ID   Q93N56      PRELIMINARY;      PRT;   547 AA.
AC   Q93N56;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 62.8 KDA PROTEIN.
OS   Coxiella burnetii.
OC   Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC   Coxiella group; Coxiella.
OX   NCBI_TaxID=777;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hoover T.A., Vodkin M.H., Williams J.C., Culp D.W., Thompson H.A.;
RT   "A chromosomal DNA deletion explains the phenotype of the Coxiella
RT   burnetii phase II variant.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF387640; AAK71266.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   547 AA;  62831 MW;  FACC03057E3CE23B CRC64;


  Query Match            100.0%;  Score 26;  DB 2;  Length 547;
  Best Local Similarity  100.0%;  Pred. No. 2.6e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    366 KLLLLK 371


RESULT  14
Q9TWB0
ID   Q9TWB0      PRELIMINARY;      PRT;   560 AA.
AC   Q9TWB0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DNA POLYMERASE ALPHA, POL ALPHA.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107655; PubMed=1762904;
RA   Ridley R.G., White J.H., McAleese S.M., Goman M., Alano P.,
RA   de Vries E., Kilbey B.J.;
RT   "DNA polymerase delta: gene sequences from Plasmodium falciparum
RT   indicate that this enzyme is more highly conserved than DNA polymerase
RT   alpha.";
RL   Nucleic Acids Res. 19:6731-6736(1991).
CC   -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
CC       PYROPHOSPHATE + DNA(N).
CC   -!- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
CC       ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR
CC       DIFFERENT REACTIONS OF DNA SYNTHESIS (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR   InterPro; IPR002064; DNA_pol_B.
DR   InterPro; IPR000209; Peptidase_S8.
DR   Pfam; PF00136; DNA_pol_B; 1.
DR   Pfam; PF03104; DNA_pol_B_exo; 1.
DR   PRINTS; PR00106; DNAPOLB.
DR   SMART; SM00486; POLBc; 1.
DR   PROSITE; PS00116; DNA_POLYMERASE_B; 1.
DR   PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
KW   DNA replication; DNA-binding; DNA-directed DNA polymerase.
SQ   SEQUENCE   560 AA;  65498 MW;  5A182E72761358D9 CRC64;


  Query Match            100.0%;  Score 26;  DB 5;  Length 560;
  Best Local Similarity  100.0%;  Pred. No. 2.7e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db    482 KLLLLK 487


RESULT  15
Q95TP6
ID   Q95TP6      PRELIMINARY;      PRT;   576 AA.
AC   Q95TP6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LD32549P.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Y, CN BW SP;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY058628; AAL13857.1; -.
SQ   SEQUENCE   576 AA;  63439 MW;  0588C7F8E07E1249 CRC64;


  Query Match            100.0%;  Score 26;  DB 5;  Length 576;
  Best Local Similarity  100.0%;  Pred. No. 2.7e+02;
```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 KLLLLK 6
           ||||||
Db     347 KLLLLK 352
```

Search completed: June 17, 2002, 12:44:20
Job time: 295 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:42:03 ; Search time 34.71 Seconds
(without alignments)
4.222 Million cell updates/sec

Title: US-09-367-714A-28
Perfect score: 26
Sequence: 1 KLLLLK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	100.0	8	3	US-08-881-971-4	Sequence 4, Appli
2	26	100.0	8	3	US-08-881-971-6	Sequence 6, Appli
3	26	100.0	13	3	US-08-881-971-5	Sequence 5, Appli
4	26	100.0	18	3	US-08-881-971-3	Sequence 3, Appli
5	26	100.0	21	1	US-07-715-397A-1	Sequence 1, Appli
6	26	100.0	21	1	US-08-060-833-1	Sequence 1, Appli
7	26	100.0	21	1	US-08-735-171-1	Sequence 1, Appli
8	26	100.0	21	1	US-08-419-824-1	Sequence 1, Appli
9	26	100.0	21	2	US-08-826-261-1	Sequence 1, Appli
10	26	100.0	21	3	US-08-848-580-1	Sequence 1, Appli
11	26	100.0	21	3	US-08-881-971-1	Sequence 1, Appli
12	26	100.0	21	5	PCT-US92-04537-7	Sequence 7, Appli
13	26	100.0	2431	1	US-07-920-281C-2	Sequence 2, Appli
14	26	100.0	2431	4	US-08-466-277-2	Sequence 2, Appli
15	24	92.3	187	4	US-09-088-549-1	Sequence 1, Appli
16	24	92.3	401	2	US-08-596-111B-2	Sequence 2, Appli
17	24	92.3	401	4	US-09-434-774-10	Sequence 10, Appl
18	24	92.3	564	3	US-08-425-843-8	Sequence 8, Appli
19	24	92.3	565	3	US-08-425-843-3	Sequence 3, Appli
20	24	92.3	905	4	US-09-360-186-3	Sequence 3, Appli
21	24	92.3	1066	4	US-09-541-782-8	Sequence 8, Appli
22	24	92.3	2052	3	US-09-045-201A-2	Sequence 2, Appli
23	23	88.5	1226	2	US-08-540-804-12	Sequence 12, Appl
24	23	88.5	1226	2	US-08-218-265-12	Sequence 12, Appl
25	23	88.5	1226	3	US-08-521-872-12	Sequence 12, Appl
26	23	88.5	1226	4	US-08-590-399-12	Sequence 12, Appl
27	22	84.6	9	2	US-08-621-803-219	Sequence 219, App
28	22	84.6	9	2	US-08-621-803-223	Sequence 223, App
29	22	84.6	9	2	US-08-621-259A-211	Sequence 211, App
30	22	84.6	9	2	US-08-621-259A-215	Sequence 215, App
31	22	84.6	9	4	US-09-217-352-219	Sequence 219, App
32	22	84.6	9	4	US-09-217-352-223	Sequence 223, App
33	22	84.6	46	2	US-08-312-202B-2	Sequence 2, Appli
34	22	84.6	46	3	US-09-079-347-2	Sequence 2, Appli
35	22	84.6	46	3	US-09-075-725-2	Sequence 2, Appli
36	22	84.6	46	4	US-08-809-646-2	Sequence 2, Appli
37	22	84.6	46	5	PCT-US95-12433-2	Sequence 2, Appli
38	22	84.6	100	4	US-09-177-249-246	Sequence 246, App
39	22	84.6	120	4	US-09-173-151A-12	Sequence 12, Appl
40	22	84.6	160	4	US-09-173-151A-6	Sequence 6, Appli
41	22	84.6	309	1	US-08-236-918A-2	Sequence 2, Appli
42	22	84.6	309	4	US-09-150-864A-2	Sequence 2, Appli
43	22	84.6	365	1	US-08-674-612-2	Sequence 2, Appli
44	22	84.6	365	2	US-08-920-296-2	Sequence 2, Appli
45	22	84.6	365	2	US-08-746-788-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT   1
US-08-881-971-4
; Sequence 4, Application US/08881971
; Patent No. 6013764
;  GENERAL INFORMATION:
;    APPLICANT:  Abdel-Magid, Ahmed F.
;    APPLICANT:  Eggmann, Urs
;    APPLICANT:  Maryanoff, Cynthia A.
;    APPLICANT:  Thaler, Adrian
;    APPLICANT:  Villani, Frank J.
;    TITLE OF INVENTION:  LIQUID PHASE PEPTIDE SYNTHESIS OF KL-4
;    TITLE OF INVENTION:  PULMONARY SURFACTANT PROTEIN
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Johnson & Johnson
;      STREET:  One Johnson & Johnson Plaza
;      CITY:  New Brunswick
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08933-003
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/881,971
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/021,455
;      FILING DATE:  17-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Kenneth J.
;      REGISTRATION NUMBER:  32,890
;      REFERENCE/DOCKET NUMBER:  MCN-586
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-524-2641
;      TELEFAX:  908-524-2808
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-881-971-4


  Query Match             100.0%;  Score 26;  DB 3;  Length 8;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db       1 KLLLLK 6


RESULT   2
US-08-881-971-6
; Sequence 6, Application US/08881971
; Patent No. 6013764
;  GENERAL INFORMATION:
;    APPLICANT:  Abdel-Magid, Ahmed F.
;    APPLICANT:  Eggmann, Urs
;    APPLICANT:  Maryanoff, Cynthia A.
;    APPLICANT:  Thaler, Adrian
;    APPLICANT:  Villani, Frank J.
;    TITLE OF INVENTION:  LIQUID PHASE PEPTIDE SYNTHESIS OF KL-4
;    TITLE OF INVENTION:  PULMONARY SURFACTANT PROTEIN
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Johnson & Johnson
;      STREET:  One Johnson & Johnson Plaza
;      CITY:  New Brunswick
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08933-003
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/881,971
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/021,455
;      FILING DATE:  17-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Kenneth J.
;      REGISTRATION NUMBER:  32,890
;      REFERENCE/DOCKET NUMBER:  MCN-586
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-524-2641
;      TELEFAX:  908-524-2808
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-881-971-6


  Query Match             100.0%;  Score 26;  DB 3;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db       3 KLLLLK 8


RESULT   3
US-08-881-971-5
; Sequence 5, Application US/08881971
; Patent No. 6013764
;  GENERAL INFORMATION:
;    APPLICANT:  Abdel-Magid, Ahmed F.
;    APPLICANT:  Eggmann, Urs
;    APPLICANT:  Maryanoff, Cynthia A.
;    APPLICANT:  Thaler, Adrian
;    APPLICANT:  Villani, Frank J.
;    TITLE OF INVENTION:  LIQUID PHASE PEPTIDE SYNTHESIS OF KL-4
;    TITLE OF INVENTION:  PULMONARY SURFACTANT PROTEIN
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Johnson & Johnson
;      STREET:  One Johnson & Johnson Plaza
;      CITY:  New Brunswick
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08933-003
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/881,971
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/021,455
;      FILING DATE:  17-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Kenneth J.
;      REGISTRATION NUMBER:  32,890
;      REFERENCE/DOCKET NUMBER:  MCN-586
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-524-2641
;      TELEFAX:  908-524-2808
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-881-971-5


  Query Match             100.0%;  Score 26;  DB 3;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db       3 KLLLLK 8


RESULT   4
US-08-881-971-3
; Sequence 3, Application US/08881971
; Patent No. 6013764
;  GENERAL INFORMATION:
;    APPLICANT:  Abdel-Magid, Ahmed F.
;    APPLICANT:  Eggmann, Urs
;    APPLICANT:  Maryanoff, Cynthia A.
;    APPLICANT:  Thaler, Adrian
;    APPLICANT:  Villani, Frank J.
;    TITLE OF INVENTION:  LIQUID PHASE PEPTIDE SYNTHESIS OF KL-4
;    TITLE OF INVENTION:  PULMONARY SURFACTANT PROTEIN
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Johnson & Johnson
;      STREET:  One Johnson & Johnson Plaza
;      CITY:  New Brunswick
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08933-003
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/881,971
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/021,455
;      FILING DATE:  17-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Kenneth J.
;      REGISTRATION NUMBER:  32,890
;      REFERENCE/DOCKET NUMBER:  MCN-586
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-524-2641
;      TELEFAX:  908-524-2808
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-881-971-3


  Query Match             100.0%;  Score 26;  DB 3;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 6.4;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db      3 KLLLLK 8


RESULT   5
US-07-715-397A-1
; Sequence 1, Application US/07715397A
; Patent No. 5260273
;  GENERAL INFORMATION:
;    APPLICANT:  Cochran, Charles G
;    APPLICANT:  Revak, Susan D
;    TITLE OF INVENTION:  PULMONARY SURFACTANT PROTEIN AND RELATED
;    TITLE OF INVENTION:  POLYPEPTIDE
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Scripps Research Institute, Office of
;      ADDRESSEE:  Patent Counsel
;      STREET:  3366 No. 5260273th Torrey Pines Ct., Suite 240
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/715,397A
;      FILING DATE:  19910614
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  SCR0395P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  internal
US-07-715-397A-1


  Query Match             100.0%;  Score 26;  DB 1;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db      1 KLLLLK 6


RESULT   6
US-08-060-833-1
; Sequence 1, Application US/08060833
; Patent No. 5407914
;  GENERAL INFORMATION:
;    APPLICANT:  Cochrane, Charles G
;    APPLICANT:  Revak, Susan D
;    TITLE OF INVENTION:  PULMONARY SURFACTANT PROTEIN AND RELATED
;    TITLE OF INVENTION:  POLYPEPTIDES
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Scripps Research Institute, Office of
;      ADDRESSEE:  Patent Counsel
;      STREET:  10666 No. 5407914th Torrey Pines Road, TPC8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/060,833
;      FILING DATE:  19930512
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Logan, April C
;      REGISTRATION NUMBER:  33,950
;      REFERENCE/DOCKET NUMBER:  SCR1309P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  internal
US-08-060-833-1


  Query Match             100.0%;  Score 26;  DB 1;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db      1 KLLLLK 6
```

```
RESULT   7
US-08-735-171-1
; Sequence 1, Application US/08735171
; Patent No. 5741891
;  GENERAL INFORMATION:
;    APPLICANT:  Weber, James V.
;    APPLICANT:  Kasulanis, Charles F.
;    APPLICANT:  Sampino, Keith
;    TITLE OF INVENTION:  PULMONARY SURFACTANT PEPTIDE
;    TITLE OF INVENTION:  SOLUBILIZATION PROCESS
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Johnson & Johnson
;      STREET:  One Johnson & Johnson Plaza
;      CITY:  New Brunswick
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08933-003
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/735,171
;      FILING DATE:  22-OCT-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Kenneth J.
;      REGISTRATION NUMBER:  32,890
;      REFERENCE/DOCKET NUMBER:  ORT-0812
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-524-2641
;      TELEFAX:  908-524-2808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-735-171-1


  Query Match             100.0%;  Score 26;  DB 1;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db       1 KLLLLK 6


RESULT   8
US-08-419-824-1
; Sequence 1, Application US/08419824
; Patent No. 5789381
;  GENERAL INFORMATION:
;    APPLICANT:  Cochrane, Charles G
;    APPLICANT:  Revak, Susan D
;    TITLE OF INVENTION:  PULMONARY SURFACTANT PROTEIN AND RELATED
;    TITLE OF INVENTION:  POLYPEPTIDES
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Scripps Research Institute, Office of
;      ADDRESSEE:  Patent Counsel
;      STREET:  10666 No. 5789381th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/419,824
;      FILING DATE:  11-APR-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/060,833
;      FILING DATE:  12-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Logan, April C
;      REGISTRATION NUMBER:  33,950
;      REFERENCE/DOCKET NUMBER:  TSRI 147.2CON2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-419-824-1


  Query Match             100.0%;  Score 26;  DB 1;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLLK 6
           ||||||
Db       1 KLLLLK 6


RESULT   9
US-08-826-261-1
; Sequence 1, Application US/08826261
; Patent No. 5952303
;  GENERAL INFORMATION:
;    APPLICANT:  Bornstein, Michael
;    APPLICANT:  Williams, N. Adeyinka
;    TITLE OF INVENTION:  LYOPHILIZED PULMONARY SURFACTANT PEPTIDE COMPOSITIONS
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Johnson & Johnson
;      STREET:  One Johnson & Johnson Plaza
;      CITY:  New Brunswick
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08933-003
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/826,261
;      FILING DATE:  6-March-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Kenneth J.
;      REGISTRATION NUMBER:  32,890
;      REFERENCE/DOCKET NUMBER:  ORT-0822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-524-2641
;      TELEFAX:  732-524-2808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  peptide
US-08-826-261-1


  Query Match             100.0%;  Score 26;  DB 2;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db      1 KLLLLK 6


RESULT  10
US-08-848-580-1
; Sequence 1, Application US/08848580
; Patent No. 6013619
;  GENERAL INFORMATION:
;    APPLICANT:  Cochrane, Charles G
;    APPLICANT:  Revak, Susan D
;    TITLE OF INVENTION:  NOVEL PULMONARY SURFACTANTS AND
;    TITLE OF INVENTION:  THERAPEUTIC USES, INCLUDING PULMONARY LAVAGE
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THE SCRIPPS RESEARCH INSTITUTE
;      STREET:  10550 No. 6013619th Torrey Pines Road, TPC-8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/848,580
;      FILING DATE:  28-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/488,123
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/419,824
;      FILING DATE:  11-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/060,833
;      FILING DATE:  12-MAY-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/715,397
;      FILING DATE:  14-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/293,201
;      FILING DATE:  04-JAN-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/141,200
;      FILING DATE:  06-JAN-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  TSRI 147.5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-784-2937
;      TELEFAX:  619-784-9399
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-848-580-1


  Query Match             100.0%;  Score 26;  DB 3;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db      1 KLLLLK 6


RESULT  11
US-08-881-971-1
; Sequence 1, Application US/08881971
; Patent No. 6013764
;  GENERAL INFORMATION:
;    APPLICANT:  Abdel-Magid, Ahmed F.
;    APPLICANT:  Eggmann, Urs
;    APPLICANT:  Maryanoff, Cynthia A.
;    APPLICANT:  Thaler, Adrian
;    APPLICANT:  Villani, Frank J.
;    TITLE OF INVENTION:  LIQUID PHASE PEPTIDE SYNTHESIS OF KL-4
;    TITLE OF INVENTION:  PULMONARY SURFACTANT PROTEIN
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Johnson & Johnson
;      STREET:  One Johnson & Johnson Plaza
;      CITY:  New Brunswick
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08933-003
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/881,971
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/021,455
;      FILING DATE:  17-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Kenneth J.
;      REGISTRATION NUMBER:  32,890
;      REFERENCE/DOCKET NUMBER:  MCN-586
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-524-2641
;      TELEFAX:  908-524-2808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-881-971-1


  Query Match             100.0%;  Score 26;  DB 3;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db      1 KLLLLK 6


RESULT  12
PCT-US92-04537-7
; Sequence 7, Application PC/TUS9204537
;  GENERAL INFORMATION:
;    APPLICANT:  Cochrane, Charles G
;    APPLICANT:  Revak, Susan D
```

```
;    TITLE OF INVENTION:  SYNTHETIC PULMONARY SURFACTANT PEPTIDES
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Scripps Research Institute, Office of
;      ADDRESSEE:  Patent Counsel
;      STREET:  10666 North Torrey Pines Road, Mail Drop TPC8
;      CITY:  La Jolla
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/04537
;      FILING DATE:  19920601
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/715,397
;      FILING DATE:  14-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bingham, Douglas A
;      REGISTRATION NUMBER:  32,457
;      REFERENCE/DOCKET NUMBER:  SCR1025P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US92-04537-7


  Query Match             100.0%;  Score 26;  DB 5;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db      1 KLLLLK 6


RESULT  13
US-07-920-281C-2
; Sequence 2, Application US/07920281C
; Patent No. 5739026
;  GENERAL INFORMATION:
;    APPLICANT:  Garoff, Henrik
;    APPLICANT:  Liljestrom, Peter
;    TITLE OF INVENTION:  DNA Expression Systems Based on
;    TITLE OF INVENTION:  Alphaviruses
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  P.O. Box 747
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/920,281C
;      FILING DATE:  13-AUG-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  828-103P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2431 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-920-281C-2


  Query Match             100.0%;  Score 26;  DB 1;  Length 2431;
  Best Local Similarity   100.0%;  Pred. No. 8.6e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db   1863 KLLLLK 1868


RESULT  14
US-08-466-277-2
; Sequence 2, Application US/08466277
; Patent No. 6190666
;  GENERAL INFORMATION:
;       APPLICANT: Garoff, Henrik
;                  Liljestrom, Peter
;       TITLE OF INVENTION: DNA Expression Systems Based on
;                           Alphaviruses
;       NUMBER OF SEQUENCES: 27
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Birch, Stewart, Kolasch & Birch
;            STREET: P.O. Box 747
;            CITY: Falls Church
;            STATE: Virginia
;            COUNTRY: USA
;            ZIP: 22040-0747
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/466,277
;            FILING DATE: 06-Jun-1995
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 07/920,281
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Murphy Jr., Gerald M.
;            REGISTRATION NUMBER: 28,977
;            REFERENCE/DOCKET NUMBER: 828-103P
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-241-1300
;            TELEFAX: 703-241-2848
;            TELEX: 248345
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2431 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-466-277-2
```

```
  Query Match            100.0%;  Score 26;  DB 4;  Length 2431;
  Best Local Similarity  100.0%;  Pred. No. 8.6e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||||||
Db   1863 KLLLLK 1868


RESULT  15
US-09-088-549-1
; Sequence 1, Application US/09088549
; Patent No. 6231853
;  GENERAL INFORMATION:
;    APPLICANT:  HILLMAN, JENNIFER L.
;    APPLICANT:  CORLEY, NEIL C.
;    APPLICANT:  PATTERSON, CHANDRA
;    TITLE OF INVENTION:  HUMAN GLUTATHIONE PEROXIDASE-6
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/088,549
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cerrone, Michael C
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0530 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-855-0572
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  187 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PROSNOT20
;      CLONE:  1817518
US-09-088-549-1


  Query Match            92.3%;  Score 24;  DB 4;  Length 187;
  Best Local Similarity  83.3%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLLK 6
          ||:|||
Db    178 KLILLK 183


Search completed: June 17, 2002, 12:42:04
Job time: 224 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:41:22 ; Search time 94.14 Seconds
(without alignments)
9.439 Million cell updates/sec

Title: US-09-367-714A-29
Perfect score: 35
Sequence: 1 KLLLKLLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	35	100.0	8	18	AAW35155	Leu/Lys diastereom
2	35	100.0	8	19	AAW82853	Antipathogenic pep
3	35	100.0	8	21	AAB17419	Antipathogenic pep
4	35	100.0	12	18	AAW35149	Leu/Lys diastereom
5	35	100.0	12	18	AAW35152	Leu/Lys diastereom
6	35	100.0	12	19	AAW82847	Antipathogenic pep
7	35	100.0	12	19	AAW82850	Antipathogenic pep
8	35	100.0	12	19	AAW82856	Antipathogenic pep
9	35	100.0	12	21	AAB17413	Antipathogenic pep
10	35	100.0	12	21	AAB17416	Antipathogenic pep
11	35	100.0	12	21	AAB17483	Antipathogenic pep
12	35	100.0	12	21	AAB17485	Antipathogenic pep
13	35	100.0	13	18	AAW35231	Diastereomer pepti
14	35	100.0	13	21	AAB17482	Antipathogenic pep
15	35	100.0	14	19	AAW82854	Antipathogenic pep
16	35	100.0	37	19	AAW77378	Lytic peptide with
17	35	100.0	77	19	AAW82858	Antipathogenic pep
18	35	100.0	77	19	AAW82859	Antipathogenic pep
19	33	94.3	18	13	AAR20977	Sequence of amphip
20	33	94.3	18	13	AAR22830	Amphiphilic peptid
21	31	88.6	153	20	AAY29393	Sperm whale myoglo
22	31	88.6	359	20	AAW81359	Human alpha-2-3 si
23	31	88.6	414	22	AAE10702	Mouse GM3 synthase
24	30	85.7	11	22	AAB97447	Peptide nucleic ac
25	30	85.7	15	19	AAW77384	Lytic peptide with
26	30	85.7	21	19	AAW62968	Minimalist lytic p
27	30	85.7	21	22	AAU03187	Membrane active sy
28	30	85.7	21	22	AAB60066	KL3 membrane activ
29	30	85.7	74	22	ABG29957	Novel human diagno
30	30	85.7	125	22	AAU17004	Human novel secret
31	30	85.7	135	22	AAY97766	I. scapularis Salp
32	30	85.7	205	22	ABG29958	Novel human diagno
33	30	85.7	216	21	AAB08846	A human MDMIP-bind
34	30	85.7	283	22	AAB60324	Helicobacter pylor
35	30	85.7	284	16	AAR75397	Human double minut
36	30	85.7	284	16	AAR75494	Human double minut
37	30	85.7	295	18	AAW20723	H. pylori cytoplas
38	30	85.7	314	21	AAB43933	Human cancer assoc
39	30	85.7	314	22	AAE03235	Human gene 12 enco
40	30	85.7	315	18	AAW19946	Alzheimer's diseas
41	30	85.7	361	22	ABG29956	Novel human diagno
42	30	85.7	377	22	AAU16931	Human novel secret
43	30	85.7	413	22	ABG21008	Novel human diagno
44	30	85.7	489	14	AAR42176	Murine MDM2. Mus
45	30	85.7	489	16	AAR76697	Mouse MDM2 protein

ALIGNMENTS

RESULT 1
AAW35155
ID AAW35155 standard; peptide; 8 AA.
XX
AC AAW35155;
XX
DT 14-APR-1998 (first entry)
XX
DE Leu/Lys diastereomer peptide [D]-L2,4,6-K3L5.
XX
KW Leu/Lys diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW Honey bee venom; pardaxin; cytolytic activity; cancer;
KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW agricultural pesticide; cell wall lysis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "D-form residue"
FT Misc-difference 4
FT /note= "D-form residue"
FT Misc-difference 6
FT /note= "D-form residue"
XX
PN WO9731019-A2.
XX
PD 28-AUG-1997.
XX
PF 20-FEB-1997; 97WO-IL00066.
XX
PR 22-FEB-1996; 96IL-0117223.
XX
PA (YEDA) YEDA RES & DEV CO LTD.

XX
PI Oren Z, Shai Y;
XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer
XX
PS Claim 21; Page 40; 80pp; English.
XX
CC This sequence represents a Leu/Lys diastereomer peptide of the
CC invention. The peptides of the invention have: (a) cytolytic activity on
CC pathogenic cells (pathogens and malignant cells not naturally present in
CC the body); but (b) no haemolytic activity, or such activity only at a
CC concentration significantly higher than that at which they lyse
CC pathogens. The peptides, their complexes and mixtures are used to treat
CC infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC or cancer, in human and veterinary medicine. Also, they can be used as
CC preservatives for food, cosmetics and agricultural produce, or as
CC agricultural pesticides. The absence of haemolytic activity (associated
CC with disturbance of alpha-helical structures) means that the peptides
CC have few if any toxic effects, and those that include D-aa will have
CC increased resistance to proteolytic degradation. Non-haemolytic,
CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC activity, allowing selection of agents for particular applications. Since
CC these random copolymers induce total lysis of bacterial cell walls,
CC resistance to them is unlikely to develop.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 35; DB 18; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 1 klllkllk 8

RESULT 2
AAW82853
ID AAW82853 standard; peptide; 8 AA.
XX
AC AAW82853;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 13; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 35; DB 19; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 1 klllkllk 8

RESULT 3
AAB17419
ID AAB17419 standard; Peptide; 8 AA.
XX
AC AAB17419;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:523.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and

```
PT   autoimmune diseases -
XX
PS   Claim 39; Page 380; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b,  where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC   -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC   where P1, P2, P3, and P4 = are each independently sequences of
CC   pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC   independently linkers; and a, b, c, d, e, and f = are each independently
CC   0 or 1, provided that at least 1 of a and b is 1. The composition can
CC   have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC   activities. DNAs, vectors and host cells from the present invention can
CC   be used for producing pharmaceutical compositions. The compositions are
CC   useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC   The use of an Fc domain (rather than a Fab domain) can provide a longer
CC   half-life or incorporate functions such as Fc receptor binding, protein
CC   A binding, complement fixation, and possibly placental transfer. AAA69443
CC   to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   8 AA;


  Query Match             100.0%;  Score 35;  DB 21;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
           ||||||||
Db       1 klllkllk 8


RESULT   4
AAW35149
ID   AAW35149 standard; peptide; 12 AA.
XX
AC   AAW35149;
XX
DT   14-APR-1998  (first entry)
XX
DE   Leu/Lys diastereomer peptide [D]-L3,4,8,10-K4L8.
XX
KW   Leu/Lys diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW   Honey bee venom; pardaxin; cytolytic activity; cancer;
KW   non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW   agricultural pesticide; cell wall lysis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 3
FT                   /note= "D-form residue"
FT   Misc-difference 4
FT                   /note= "D-form residue"
FT   Misc-difference 8
FT                   /note= "D-form residue"
FT   Misc-difference 10
FT                   /note= "D-form residue"
FT   Modified-site   12
FT                   /note= "C-terminal amide"
XX
PN   WO9731019-A2.
XX
PD   28-AUG-1997.
XX
PF   20-FEB-1997;   97WO-IL00066.
XX
PR   22-FEB-1996;   96IL-0117223.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Oren Z,  Shai Y;
XX
DR   WPI; 1997-435088/40.
XX
PT   Peptide(s) having selective cytolytic activity - against pathogens
PT   and malignant cells, but no haemolytic activity, used for treating
PT   infections and cancer
XX
PS   Claim 21; Page 39; 80pp; English.
XX
CC   This sequence represents a Leu/Lys diastereomer peptide of the
CC   invention. The peptides of the invention have: (a) cytolytic activity on
CC   pathogenic cells (pathogens and malignant cells not naturally present in
CC   the body); but (b) no haemolytic activity, or such activity only at a
CC   concentration significantly higher than that at which they lyse
CC   pathogens. The peptides, their complexes and mixtures are used to treat
CC   infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC   or cancer, in human and veterinary medicine. Also, they can be used as
CC   preservatives for food, cosmetics and agricultural produce, or as
CC   agricultural pesticides. The absence of haemolytic activity (associated
CC   with disturbance of alpha-helical structures) means that the peptides
CC   have few if any toxic effects, and those that include D-aa will have
CC   increased resistance to proteolytic degradation. Non-haemolytic,
CC   cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC   activity, allowing selection of agents for particular applications. Since
CC   these random copolymers induce total lysis of bacterial cell walls,
CC   resistance to them is unlikely to develop.
XX
SQ   Sequence   12 AA;


  Query Match             100.0%;  Score 35;  DB 18;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 2.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
           ||||||||
Db       5 klllkllk 12


RESULT   5
AAW35152
ID   AAW35152 standard; peptide; 12 AA.
XX
AC   AAW35152;
XX
DT   14-APR-1998  (first entry)
XX
DE   Leu/Lys diastereomer peptide [D]-K1,5,9,12L2,6,7,11-K4L8.
XX
KW   Leu/Lys diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW   Honey bee venom; pardaxin; cytolytic activity; cancer;
KW   non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW   agricultural pesticide; cell wall lysis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 2
FT                   /note= "D-form residue"
FT   Misc-difference 5
FT                   /note= "D-form residue"
FT   Misc-difference 6
FT                   /note= "D-form residue"
FT   Misc-difference 7
FT                   /note= "D-form residue"
FT   Misc-difference 9
FT                   /note= "D-form residue"
FT   Misc-difference 11
FT                   /note= "D-form residue"
```

FT Misc-difference 12
FT /note= "D-form residue"
FT Modified-site 12
FT /note= "C-terminal amide"
XX
PN WO9731019-A2.
XX
PD 28-AUG-1997.
XX
PF 20-FEB-1997; 97WO-IL00066.
XX
PR 22-FEB-1996; 96IL-0117223.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer
XX
PS Claim 21; Page 40; 80pp; English.
XX
CC This sequence represents a Leu/Lys diastereomer peptide of the
CC invention. The peptides of the invention have: (a) cytolytic activity on
CC pathogenic cells (pathogens and malignant cells not naturally present in
CC the body); but (b) no haemolytic activity, or such activity only at a
CC concentration significantly higher than that at which they lyse
CC pathogens. The peptides, their complexes and mixtures are used to treat
CC infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC or cancer, in human and veterinary medicine. Also, they can be used as
CC preservatives for food, cosmetics and agricultural produce, or as
CC agricultural pesticides. The absence of haemolytic activity (associated
CC with disturbance of alpha-helical structures) means that the peptides
CC have few if any toxic effects, and those that include D-aa will have
CC increased resistance to proteolytic degradation. Non-haemolytic,
CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC activity, allowing selection of agents for particular applications. Since
CC these random copolymers induce total lysis of bacterial cell walls,
CC resistance to them is unlikely to develop.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 18; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 5 klllkllk 12

RESULT 6
AAW82847
ID AAW82847 standard; peptide; 12 AA.
XX
AC AAW82847;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 12; Page 105; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 5 klllkllk 12

RESULT 7
AAW82850
ID AAW82850 standard; peptide; 12 AA.
XX
AC AAW82850;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.

XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 13; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 5 klllkllk 12

RESULT 8
AAW82856
ID AAW82856 standard; peptide; 12 AA.
XX
AC AAW82856;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 14; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 5 klllkllk 12

RESULT 9
AAB17413
ID AAB17413 standard; Peptide; 12 AA.
XX
AC AAB17413;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:517.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX

DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 378; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 5 klllkllk 12

RESULT 10
AAB17416
ID AAB17416 standard; Peptide; 12 AA.
XX
AC AAB17416;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:520.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.

XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 379; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 5 klllkllk 12

RESULT 11
AAB17483
ID AAB17483 standard; Peptide; 12 AA.
XX
AC AAB17483;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:587.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX

```
PT   Novel composition of matter comprising an Fc domain and
PT   pharmacologically active peptides, useful for treating cancer and
PT   autoimmune diseases -
XX
PS   Claim 39; Page 401; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b,  where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC   -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC   where P1, P2, P3, and P4 = are each independently sequences of
CC   pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC   independently linkers; and a, b, c, d, e, and f = are each independently
CC   0 or 1, provided that at least 1 of a and b is 1. The composition can
CC   have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC   activities. DNAs, vectors and host cells from the present invention can
CC   be used for producing pharmaceutical compositions. The compositions are
CC   useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC   The use of an Fc domain (rather than a Fab domain) can provide a longer
CC   half-life or incorporate functions such as Fc receptor binding, protein
CC   A binding, complement fixation, and possibly placental transfer. AAA69443
CC   to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   12 AA;


  Query Match             100.0%;  Score 35;  DB 21;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 2.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          ||||||||
Db      5 klllkllk 12


RESULT  12
AAB17485
ID   AAB17485 standard; Peptide; 12 AA.
XX
AC   AAB17485;
XX
DT   31-OCT-2000  (first entry)
XX
DE   Antipathogenic peptide sequence SEQ ID NO:589.
XX
KW   Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW   autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW   immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW   MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW   cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW   vascular endothelial growth factor; matrix metalloproteinase;
KW   asthma; thrombosis; pharmaceutical.
XX
OS   Synthetic.
XX
PN   WO200024782-A2.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-US25044.
XX
PR   23-OCT-1998;   98US-0105371.
PR   22-OCT-1999;   99US-0428082.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham J,  Boone TC;
XX
DR   WPI; 2000-350702/30.
XX
PT   Novel composition of matter comprising an Fc domain and
PT   pharmacologically active peptides, useful for treating cancer and
PT   autoimmune diseases -
XX
PS   Claim 39; Page 402; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b,  where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC   -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC   where P1, P2, P3, and P4 = are each independently sequences of
CC   pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC   independently linkers; and a, b, c, d, e, and f = are each independently
CC   0 or 1, provided that at least 1 of a and b is 1. The composition can
CC   have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC   activities. DNAs, vectors and host cells from the present invention can
CC   be used for producing pharmaceutical compositions. The compositions are
CC   useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC   The use of an Fc domain (rather than a Fab domain) can provide a longer
CC   half-life or incorporate functions such as Fc receptor binding, protein
CC   A binding, complement fixation, and possibly placental transfer. AAA69443
CC   to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   12 AA;


  Query Match             100.0%;  Score 35;  DB 21;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 2.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          ||||||||
Db      5 klllkllk 12


RESULT  13
AAW35231
ID   AAW35231 standard; peptide; 13 AA.
XX
AC   AAW35231;
XX
DT   14-APR-1998  (first entry)
XX
DE   Diastereomer peptide [D]-L3,4,8,10-K4L8C.
XX
KW   Diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW   Honey bee venom; pardaxin; cytolytic activity; cancer;
KW   non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW   agricultural pesticide; cell wall lysis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 3
FT                   /note= "D-form residue"
FT   Misc-difference 4
FT                   /note= "D-form residue"
FT   Misc-difference 8
FT                   /note= "D-form residue"
FT   Misc-difference 10
FT                   /note= "D-form residue"
XX
PN   WO9731019-A2.
XX
PD   28-AUG-1997.
XX
PF   20-FEB-1997;   97WO-IL00066.
XX
PR   22-FEB-1996;   96IL-0117223.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
```

PI Oren Z, Shai Y;
XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer
XX
PS Example 7; Page 49; 80pp; English.
XX
CC This sequence represents a diastereomer peptide of the invention. This
CC sequence is used in a "bundle sequence", which is created by binding 5
CC copies of this sequence to peptide 23 (see AAW35149). The peptides of
CC the invention have: (a) cytolytic activity on pathogenic cells (pathogens
CC and malignant cells not naturally present in the body); but (b) no
CC haemolytic activity, or such activity only at a concentration
CC significantly higher than that at which they lyse pathogens. The
CC peptides, their complexes and mixtures are used to treat infections
CC (caused by bacteria, fungi, protozoa, mycoplasma or viruses) or cancer,
CC in human and veterinary medicine. Also, they can be used as preservatives
CC for food, cosmetics and agricultural produce, or as agricultural
CC pesticides. The absence of haemolytic activity (associated with
CC disturbance of alpha-helical structures) means that the peptides have few
CC if any toxic effects, and those that include D-aa will have increased
CC resistance to proteolytic degradation. Non-haemolytic, cytotoxic random
CC copolymers of pardaxin, each has a specific spectrum of activity,
CC allowing selection of agents for particular applications. Since these
CC random copolymers induce total lysis of bacterial cell walls, resistance
CC to them is unlikely to develop.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 35; DB 18; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 5 klllkllk 12

RESULT 14
AAB17482
ID AAB17482 standard; Peptide; 13 AA.
XX
AC AAB17482;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:586.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX

PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 401; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 35; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||||
Db 5 klllkllk 12

RESULT 15
AAW82854
ID AAW82854 standard; peptide; 14 AA.
XX
AC AAW82854;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the

PT continuity of an alpha-helical structure
XX
PS Claim 14; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 35; DB 19; Length 14;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 KLLLKLLK 8
          ||||||||
Db      6 klllkllk 13
```

Search completed: June 17, 2002, 12:41:23
Job time: 298 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:42:59 ; Search time 46.42 Seconds
(without alignments)
16.560 Million cell updates/sec

Title: US-09-367-714A-29
Perfect score: 35
Sequence: 1 KLLLKLLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	32	91.4	238	2	E71375	probable ABC trans
2	32	91.4	302	2	S27846	hypothetical prote
3	31	88.6	109	2	S42121	RNaseP C5 chain -
4	31	88.6	296	2	G97799	hypothetical prote
5	31	88.6	387	2	JE0364	lactosylceramide a
6	31	88.6	1211	2	S68251	phospholipase C, i
7	30	85.7	166	2	H82873	hypothetical prote
8	30	85.7	191	2	F90392	hypothetical prote
9	30	85.7	282	2	G71932	hypothetical prote
10	30	85.7	321	2	F71163	probable oligopept
11	30	85.7	433	2	A69735	phage PBSX termina
12	30	85.7	489	2	S15349	mdm2 protein - mou
13	30	85.7	491	1	S24354	p53-binding protei
14	30	85.7	515	2	T39031	hypothetical prote
15	30	85.7	727	2	T47638	hypothetical prote
16	30	85.7	984	1	DJNVCP	DNA-directed DNA p
17	30	85.7	986	2	T41809	DNA polymerase orf
18	30	85.7	1052	2	T00067	hypothetical prote
19	30	85.7	2513	2	G96536	hypothetical prote
20	29	82.9	155	2	T45215	transcription regu
21	29	82.9	181	2	C69057	conserved bacitrac
22	29	82.9	311	2	A88601	protein Y49E10.8 [
23	29	82.9	346	2	T51377	hypothetical prote
24	29	82.9	362	2	F90441	ABC transporter, A
25	29	82.9	384	2	G81436	probable serine/th
26	29	82.9	464	2	T48984	hypothetical prote
27	29	82.9	465	2	T30155	hypothetical prote
28	29	82.9	473	2	F70513	probable PPE prote
29	29	82.9	632	2	AF1189	transcription anti
30	29	82.9	632	2	AG1547	transcription anti
31	29	82.9	641	2	G85043	hypothetical prote
32	29	82.9	669	2	T44681	GTP-binding protei
33	29	82.9	684	2	E64496	ATP-dependent RNA
34	29	82.9	753	2	JC7386	retinovin - chicke
35	29	82.9	802	2	G89893	PriA, primosomal p
36	29	82.9	884	2	E86244	unknown protein, 4
37	29	82.9	1223	2	S62011	PHO85 protein - ye
38	29	82.9	1846	2	T10670	hypothetical prote
39	29	82.9	1941	2	T30554	ubiquitin-protein
40	28	80.0	118	2	F90459	hypothetical prote
41	28	80.0	138	2	S36115	interferon - Japan
42	28	80.0	138	2	AC1108	transcription regu
43	28	80.0	138	2	AD1469	transcription regu
44	28	80.0	149	2	C84053	sodium-glutamate/a
45	28	80.0	164	2	T03915	hypothetical prote

ALIGNMENTS

RESULT 1
E71375
probable ABC transporter, ATP-binding protein - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 17-Mar-2000
C;Accession: E71375
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770
A;Accession: E71375
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-238 <COL>
A;Cross-references: GB:AE001188; GB:AE000520; NID:g3322282; PIDN:AAC65030.1; PID:g332
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0035
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog
C;Keywords: ATP
F;27-207/Domain: ATP-binding cassette homology <ABC>

Query Match 91.4%; Score 32; DB 2; Length 238;
Best Local Similarity 87.5%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 KLLLKLLK 8
            ||:|||||
Db      55 KLVLKLLK 62
```

RESULT 2
S27846
hypothetical protein - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Sep-1997
C;Accession: S27846
R;Woodward, R.; Carden, M.J.; Gull, K.
submitted to the EMBL Data Library, March 1992
A;Reference number: S27846
A;Accession: S27846
A;Molecule type: mRNA
A;Residues: 1-302 <WOO>
A;Cross-references: EMBL:M87318; NID:g162176; PID:g162177

Query Match 91.4%; Score 32; DB 2; Length 302;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 KLLLKLLK 8
            |||||:||
Db      21 KLLLKVLK 28


RESULT   3
S42121
RNaseP C5 chain - Mycoplasma capricolum
C;Species: Mycoplasma capricolum
C;Date: 07-Sep-1994 #sequence_revision 26-May-1995 #text_change 20-Jun-2000
C;Accession: S42121
R;Miyata, M.; Sano, K.I.; Okada, R.; Fukumura, T.
Nucleic Acids Res. 21, 4816-4823, 1993
A;Title: Mapping of replication initiation site in Mycoplasma capricolum genome by two-d
A;Reference number: S42116; MUID:94051609
A;Accession: S42121
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-109 <MIY>
A;Cross-references: EMBL:D14982; NID:g416237; PIDN:BAA03619.1; PID:g416239
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: ribonuclease P, protein component

  Query Match             88.6%;  Score 31;  DB 2;  Length 109;
  Best Local Similarity   75.0%;  Pred. No. 24;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            |||:||:|
Db      99 KLLIKLIK 106


RESULT   4
G97799
hypothetical protein RC0799 [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C;Accession: G97799
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; RC
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: G97799
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-296 <KUR>
A;Cross-references: GB:AE006914; PIDN:AAL03337.1; PID:g15619897; GSPDB:GN00173
C;Genetics:
A;Gene: RC0799

  Query Match             88.6%;  Score 31;  DB 2;  Length 296;
  Best Local Similarity   75.0%;  Pred. No. 64;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            |:|||:||
Db     287 KILLKILK 294


RESULT   5
JE0364
lactosylceramide alpha-2,3-sialytransferase (EC 2.4.99.9) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C;Accession: JE0364
R;Kono, M.; Takashima, S.; Liu, H.; Inoue, M.; Kojima, N.; Lee, Y.; Hamamoto, T.; Tsuji,
Biochem. Biophys. Res. Commun. 253, 170-175, 1998
A;Title: Molecular cloning and functional expression of a fifth-type alpha2,3-sialyltran
A;Reference number: JE0364; MUID:99092398
A;Accession: JE0364
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-387 <KON>
A;Cross-references: GB:Y15003
C;Superfamily: alpha-2,3-sialyltransferase ST2
C;Keywords: glycosyltransferase

  Query Match             88.6%;  Score 31;  DB 2;  Length 387;
  Best Local Similarity   87.5%;  Pred. No. 83;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            | ||||||
Db     368 KFLLKLLK 375


RESULT   6
S68251
phospholipase C, inositol-lipid specific (EC 3.1.4.-) isoform beta - turkey
C;Species: Meleagris gallopavo (common turkey)
C;Date: 05-Dec-1996 #sequence_revision 07-Feb-1997 #text_change 22-Jun-1999
C;Accession: S68251; S72374; A61270
R;Waldo, G.L.; Paterson, A.; Boyer, J.L.; Nicholas, R.A.; Harden, T.K.
Biochem. J. 316, 559-568, 1996
A;Title: Molecular cloning, expression and regulatory activity of G-alpha(11)- and be
A;Reference number: S68251; MUID:96257751
A;Accession: S68251
A;Molecule type: mRNA
A;Residues: 1-1211 <WAL1>
A;Cross-references: GB:U49431; NID:g1223919; PIDN:AAC60011.1; PID:g1223920
A;Experimental source: erythrocyte
A;Accession: S72374
A;Molecule type: protein
A;Residues: 210-216,'M',218-231;244-248;284-291;345-353,'S',355-360;453-461;661-679 <
A;Experimental source: erythrocyte
R;Waldo, G.L.; Morris, A.J.; Klapper, D.G.; Harden, T.K.
Mol. Pharmacol. 40, 480-489, 1991
A;Title: Receptor- and G-protein-regulated 150-kDa avian phospholipase C: inhibition
ished by immunoreactivity and peptide sequence.
A;Reference number: A61270; MUID:92017673
A;Accession: A61270
A;Status: preliminary
A;Molecule type: protein
A;Residues: 284-292,'X',294-296,'R';568-577;751-753,'L',755-759;765-776,'T',778-780;8
A;Experimental source: erythrocyte
A;Note: 885-His was also found
C;Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase I; 1-phospha
odiesterase domain Y homology
C;Keywords: phosphoric diester hydrolase
F;314-463/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X
F;543-663/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain Y

  Query Match             88.6%;  Score 31;  DB 2;  Length 1211;
  Best Local Similarity   87.5%;  Pred. No. 2.5e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            || |||||
Db     902 KLFLKLLK 909


RESULT   7
H82873
hypothetical protein UU580 [imported] - Ureaplasma urealyticum
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: H82873
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
```

A;Reference number: A82870
A;Accession: H82873
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-166 <GLA>
A;Cross-references: GB:AE002156; GB:AF222894; NID:g6899580; PIDN:AAF30994.1; GSPDB:GN001
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: UU580
A;Genetic code: SGC3

Query Match 85.7%; Score 30; DB 2; Length 166;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||:|:|||
Db 60 KLILRLLK 67

RESULT 8
F90392
hypothetical protein SSO2227 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: F90392
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: F90392
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-191 <KUR>
A;Cross-references: GB:AE006641; NID:g13815527; PIDN:AAK42397.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO2227

Query Match 85.7%; Score 30; DB 2; Length 191;
Best Local Similarity 87.5%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
| ||||||
Db 36 KALLKLLK 43

RESULT 9
G71932
hypothetical protein jhp0431 - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 11-Jan-2000
C;Accession: G71932
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: A71800; MUID:99120557
A;Accession: G71932
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-282 <ARN>
A;Cross-references: GB:AE001477; GB:AE001439; NID:g4154961; PIDN:AAD06012.1; PID:g415496
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0431
C;Superfamily: Helicobacter hypothetical protein HP0479

Query Match 85.7%; Score 30; DB 2; Length 282;
Best Local Similarity 75.0%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
|::|||||
Db 171 KIVLKLLK 178

RESULT 10
F71163
probable oligopeptide transport ATP-binding protein AppF - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 17-Mar-2000
C;Accession: F71163
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophili
A;Reference number: A71000; MUID:98344137
A;Accession: F71163
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-321 <KAW>
A;Cross-references: GB:AP000002; NID:g3236129; PIDN:BAA29595.1; PID:d1030538; PID:g32
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBa
C;Genetics:
A;Gene: PH0507
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog
C;Keywords: ATP
F;27-230/Domain: ATP-binding cassette homology <ABC>

Query Match 85.7%; Score 30; DB 2; Length 321;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||:|:|||
Db 55 KLILRLLK 62

RESULT 11
A69735
phage PBSX terminase large chain xtmB - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: A69735; I40415; S47115
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033
A;Accession: A69735
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-433 <KUN>
A;Cross-references: GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB13115.1; PID:e11832
A;Experimental source: strain 168
R;McDonnell, G.E.; Wood, H.; Devine, K.M.; McConnell, D.J.

J. Bacteriol. 176, 5820-5830, 1994
A;Title: Genetic control of bacterial suicide: regulation of the induction of PBSX in Ba
A;Reference number: I40408; MUID:94364963
A;Accession: I40415
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-76 <RES>
A;Cross-references: EMBL:Z34287; NID:g498810; PIDN:CAA84048.1; PID:g498818
C;Genetics:
A;Gene: xtmB

Query Match 85.7%; Score 30; DB 2; Length 433;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
|::|||||
Db 39 KIVLKLLK 46

RESULT 12
S15349
mdm2 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Mar-2000
C;Accession: S15349
R;Fakharzadeh, S.S.; Trusko, S.P.; George, D.L.
EMBO J. 10, 1565-1569, 1991
A;Title: Tumorigenic potential associated with enhanced expression of a gene that is amp
A;Reference number: S15349; MUID:91224107
A;Accession: S15349
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-489 <FAK>
A;Cross-references: EMBL:X58876; NID:g53038; PIDN:CAA41684.1; PID:g53039
C;Genetics:
A;Gene: mdm2
C;Superfamily: human p53-binding protein mdm2

Query Match 85.7%; Score 30; DB 2; Length 489;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLKLLK 8
|||||||
Db 33 LLLKLLK 39

RESULT 13
S24354
p53-binding protein mdm2 - human
N;Alternate names: mdm-2 oncogene; mouse double minute 2 homolog; p53-associated phospho
N;Contains: p53-binding protein mdm2, splice form A
C;Species: Homo sapiens (man)
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C;Accession: S24354; S57338; G02026
R;Oliner, J.D.; Kinzler, K.W.; Meltzer, P.S.; George, D.L.; Vogelstein, B.
Nature 358, 80-83, 1992
A;Title: Amplification of a gene encoding a p53-associated protein in human sarcomas.
A;Reference number: S24354; MUID:92310576
A;Accession: S24354
A;Molecule type: mRNA
A;Residues: 1-491 <OLI>
A;Cross-references: EMBL:Z12020; NID:g35211; PIDN:CAA78055.1; PID:g35212
R;Zauberman, A.; Flusberg, D.; Haupt, Y.; Barak, Y.; Oren, M.
Nucleic Acids Res. 23, 2584-2592, 1995
A;Title: A functional p53-responsive intronic promoter is contained within the human mdm
A;Reference number: S57338; MUID:95380270
A;Accession: S57338
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-16,'P',18-24 <ZAU>
A;Cross-references: EMBL:U28935; NID:g904033; PIDN:AAA82237.1; PID:g904034
R;Lunec, J.
submitted to the EMBL Data Library, August 1995
A;Description: Multiple alternate spliced mdm2 transcripts with loss of p53 binding d
A;Reference number: G09070
A;Accession: G02026
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-27,223-491 <LUN>
A;Cross-references: EMBL:U33199; NID:g992676; PIDN:AAA75514.1; PID:g992677
A;Experimental source: splice form A
C;Genetics:
A;Gene: GDB:MDM2
A;Cross-references: GDB:250456; OMIM:164785
A;Map position: 12q14.3-12q15
C;Superfamily: human p53-binding protein mdm2
C;Keywords: alternative splicing; oncogene; phosphoprotein
F;1-491/Product: p53-binding protein mdm-2 #status predicted <MAT1>
F;1-27,223-491/Product: p53-binding protein mdm-2, splice form A #status predicted <M

Query Match 85.7%; Score 30; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLKLLK 8
|||||||
Db 33 LLLKLLK 39

RESULT 14
T39031
hypothetical protein SPAC6C3.07 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39031
R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z21750
A;Accession: T39031
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-515 <DEV>
A;Cross-references: EMBL:Z69731; PIDN:CAA93619.1; GSPDB:GN00066; SPDB:SPAC6C3.07
A;Experimental source: strain 972h-; cosmid c6C3
C;Genetics:
A;Gene: SPDB:SPAC6C3.07
A;Map position: 1

Query Match 85.7%; Score 30; DB 2; Length 515;
Best Local Similarity 87.5%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
|||||| |
Db 49 KLLLKLTK 56

RESULT 15
T47638
hypothetical protein T5N23.150 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T47638
R;Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24463
A;Accession: T47638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-727 <OBE>

A;Cross-references: EMBL:AL138650
A;Experimental source: cultivar Columbia; BAC clone T5N23
C;Genetics:
A;Map position: 3
A;Introns: 56/3; 95/3
A;Note: T5N23.150

Query Match 85.7%; Score 30; DB 2; Length 727;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 LLLKLLK 8
           |||||||
Db       3 LLLKLLK 9
```

Search completed: June 17, 2002, 12:43:01
Job time: 256 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:44:46 ; Search time 21.35 Seconds
(without alignments)
14.508 Million cell updates/sec

Title: US-09-367-714A-29
Perfect score: 35
Sequence: 1 KLLLKLLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	32	91.4	238	1	Y035_TREPA	O83078 treponema p
2	31	88.6	109	1	RNPA_MYCCA	P43039 mycoplasma
3	30	85.7	433	1	XTMB_BACSU	P39786 bacillus su
4	30	85.7	487	1	MDM2_CANFA	P56950 canis famil
5	30	85.7	489	1	MDM2_MOUSE	P23804 mus musculu
6	30	85.7	491	1	MDM2_HORSE	P56951 equus cabal
7	30	85.7	491	1	MDM2_HUMAN	Q00987 homo sapien
8	30	85.7	515	1	YD56_SCHPO	Q10310 schizosacch
9	30	85.7	723	1	GGA3_HUMAN	Q9nz52 homo sapien
10	30	85.7	984	1	DPOL_NPVAC	P18131 autographa
11	30	85.7	986	1	DPOL_NPVBM	P41712 bombyx mori
12	29	82.9	172	1	CAPG_STAAU	P39856 staphylococ
13	29	82.9	229	1	SOML_TETMU	Q9i9h4 tetraodon m
14	29	82.9	231	1	SOM1_SPAAU	P54863 sparus aura
15	29	82.9	231	1	SOM2_SPAAU	P79894 sparus aura
16	29	82.9	231	1	SOML_SCIOC	Q9ygk7 sciaenops o
17	29	82.9	231	1	SOML_SIGGU	Q9pwg4 siganus gut
18	29	82.9	1941	1	UBR1_KLULA	O60014 kluyveromyc
19	28	80.0	193	1	RL18_TRYBB	P50885 trypanosoma
20	28	80.0	367	1	IPL1_YEAST	P38991 saccharomyc
21	28	80.0	375	1	ASG2_BACSU	O34482 bacillus su
22	28	80.0	424	1	HE47_DROME	Q27268 drosophila
23	28	80.0	445	1	YPNP_BACSU	P54181 bacillus su
24	28	80.0	471	1	FTSA_STAAU	O07325 staphylococ
25	28	80.0	475	1	ASPA_BACSU	P26899 bacillus su
26	28	80.0	536	1	YC42_SYNY3	P42349 synechocyst
27	28	80.0	541	1	AAAT_RABIT	O19105 oryctolagus
28	28	80.0	553	1	AAAX_MOUSE	P51912 mus musculu
29	28	80.0	555	1	YKQC_BACSU	Q45493 bacillus su
30	28	80.0	689	1	YE30_HELPY	P56185 helicobacte
31	28	80.0	692	1	YE30_HELPJ	Q9zji6 helicobacte
32	28	80.0	1437	1	DPO3_BACSU	P13267 bacillus su
33	27	77.1	79	1	Y331_SYNY3	Q55785 synechocyst
34	27	77.1	128	1	RNPA_MYCGE	P47703 mycoplasma
35	27	77.1	180	1	YD65_SCHPO	Q10317 schizosacch
36	27	77.1	255	1	LP61_EIMTE	P15714 eimeria ten
37	27	77.1	342	1	Y755_METJA	Q58165 methanococc
38	27	77.1	344	1	ETFA_YEAST	Q12480 saccharomyc
39	27	77.1	346	1	YC89_ARCFU	O28980 archaeoglob
40	27	77.1	362	1	YAHA_ECOLI	P21514 escherichia
41	27	77.1	386	1	RMAR_HANWI	P48849 hansenula w
42	27	77.1	431	1	YQAT_BACSU	P45916 bacillus su
43	27	77.1	434	1	Y610_METJA	Q58027 methanococc
44	27	77.1	442	1	YQHB_BACSU	P54505 bacillus su
45	27	77.1	470	1	GLG1_SOLTU	Q00081 solanum tub

ALIGNMENTS

```
RESULT   1
Y035_TREPA
ID   Y035_TREPA     STANDARD;      PRT;   238 AA.
AC   O83078;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable metal transport system ATP-binding protein TP0035.
GN   TP0035.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NICHOLS;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   -!- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM
CC       TP0034/TP0035/TP0036 FOR A METAL. PROBABLY RESPONSIBLE FOR ENERGY
CC       COUPLING TO THE TRANSPORT SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE001188; AAC65030.1; -.
DR   TIGR; TP0035; -.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; FALSE_NEG.
KW   Hypothetical protein; Transport; Inner membrane; ATP-binding;
KW   Complete proteome.
FT   NP_BIND      44      51       ATP (POTENTIAL).
SQ   SEQUENCE   238 AA;  26460 MW;  673E7B4882BE4D29 CRC64;
```

Query Match 91.4%; Score 32; DB 1; Length 238;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 KLLLKLLK 8
            ||:||||
Db      55 KLVLKLLK 62


RESULT   2
RNPA_MYCCA
ID   RNPA_MYCCA      STANDARD;      PRT;   109 AA.
AC   P43039;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ribonuclease P protein component (EC 3.1.26.5) (RNaseP protein)
DE   (RNase P protein) (Protein C5).
GN   RNPA.
OS   Mycoplasma capricolum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Entomoplasmataceae.
OX   NCBI_TaxID=2095;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 27343;
RX   MEDLINE=94051609; PubMed=8233831;
RA   Miyata M., Sano K.-I., Okada R., Fukumura T.;
RT   "Mapping of replication initiation site in Mycoplasma capricolum
RT   genome by two-dimensional gel-electrophoretic analysis.";
RL   Nucleic Acids Res. 21:4816-4823(1993).
CC   -!- FUNCTION: RNaseP catalyzes the removal of the 5'-leader sequence
CC       from pre-tRNA to produce the mature 5'terminus. It can also
CC       cleave other RNA substrates such as 4.5S RNA. The protein
CC       component plays an auxiliary but essential role in vivo by binding
CC       to the 5'-leader sequence and broadening the substrate specificity
CC       of the ribozyme (By similarity).
CC   -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-
CC       extra-nucleotide from tRNA precursor.
CC   -!- SUBUNIT: Consists of a catalytic RNA component (M1 or rnpB) and a
CC       protein subunit (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RNPA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D14982; BAA03619.1; -.
DR   HSSP; P25814; 1A6F.
DR   InterPro; IPR000100; Ribonuclease_P.
DR   Pfam; PF00825; Ribonuclease_P; 1.
DR   PROSITE; PS00648; RIBONUCLEASE_P; 1.
KW   Hydrolase; Nuclease; Endonuclease; tRNA processing; RNA-binding.
SQ   SEQUENCE   109 AA;  12900 MW;  ACF520A0982C0D12 CRC64;


  Query Match             88.6%;  Score 31;  DB 1;  Length 109;
  Best Local Similarity   75.0%;  Pred. No. 8.2;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            |||:||:|
Db      99 KLLIKLIK 106


RESULT   3
XTMB_BACSU
ID   XTMB_BACSU      STANDARD;      PRT;   433 AA.
AC   P39786;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   PBSX phage terminase large subunit.
GN   XTMB.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Krogh S., O'Reilly M., Nolan N., Devine K.M.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-76 FROM N.A.
RC   STRAIN=168 / SO113;
RX   MEDLINE=94364963; PubMed=8083174;
RA   McDonnell G.E., Wood H., Devine K.M., McConnell D.J.;
RT   "Genetic control of bacterial suicide: regulation of the induction of
RT   PBSX in Bacillus subtilis.";
RL   J. Bacteriol. 176:5820-5830(1994).
CC   -!- FUNCTION: FUNCTION AS A TERMINASE.
CC   -!- SUBUNIT: DIMER OF A SMALL AND A LARGE SUBUNIT (POTENTIAL).
CC   -!- SIMILARITY: STRONG, TO B.SUBTILIS YQAT.
CC   -!- SIMILARITY: TO LARGE SUBUNIT OF B.SUBTILIS PHAGE SPP1 TERMINASE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z70177; CAA94059.1; -.
DR   EMBL; Z34287; CAA84048.1; -.
DR   EMBL; Z99110; CAB13115.1; -.
DR   PIR; S47115; S47115.
DR   SubtiList; BG11000; xtmB.
KW   DNA packaging; Complete proteome.
SQ   SEQUENCE   433 AA;  51150 MW;  471FC77DFEA2CA10 CRC64;


  Query Match             85.7%;  Score 30;  DB 1;  Length 433;
  Best Local Similarity   75.0%;  Pred. No. 51;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            |::||||
Db      39 KIVLKLLK 46


RESULT   4
MDM2_CANFA
ID   MDM2_CANFA      STANDARD;      PRT;   487 AA.
AC   P56950;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (P53-binding protein
DE   Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (Cdm2).
GN   MDM2.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE OF 1-484 FROM N.A.
RX   MEDLINE=20218866; PubMed=10754200;
RA   Nasir L., Burr P.D., McFarlane S.T., Gault E., Thompson H.,
RA   Argyle D.J.;
RT   "Cloning, sequence analysis and expression of the cDNAs encoding the
RT   canine and equine homologues of the mouse double minute 2 (mdm2)
RT   proto-oncogene.";
RL   Cancer Lett. 152:9-13(2000).
RN   [2]
```

```
RP   SEQUENCE FROM N.A. (ISOFORMS MDM2 AND MDM2-ALPHA).
RX   MEDLINE=20065171; PubMed=10597303;
RA   Veldhoen N., Metcalfe S., Milner J.;
RT   "A novel exon within the mdm2 gene modulates translation initiation in
RT   vitro and disrupts the p53-binding domain of mdm2 protein.";
RL   Oncogene 18:7026-7033(1999).
CC   -!- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
CC       APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
CC       FUNCTIONS AS AN UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
CC       TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
CC       TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS (BY SIMILARITY).
CC   -!- COFACTOR: ZINC IS REQUIRED FOR UBIQUITIN LIGASE E3 ACTIVITY (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
CC       SPECIFICALLY TO RNA. CAN INTERACTS ALSO WITH RETINOBLASTOMA
CC       PROTEIN (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1
CC       TRANSCRIPTION FACTOR (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC. EXPRESSED
CC       PREDOMINANTLY IN THE NUCLEOPLASM (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; MDM2 (SHOWN HERE) AND MDM2-
CC       ALPHA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: MDM2-ALPHA IS PRESENT IN LYMPHOID AND
CC       TESTICULAR TISSUES.
CC   -!- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
CC       G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
CC       REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
CC       INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
CC       FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
CC       ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
CC       PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
CC       ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
CC       ITSELF (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC   -!- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
CC   -!- SIMILARITY: BELONGS TO THE MDM2 / MDM4 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF100705; AAF67833.1; -.
DR   HSSP; Q9UMT8; 1YCR.
DR   InterPro; IPR003160; MDM2.
DR   InterPro; IPR001876; Znf-RanBP.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF02279; MDM2; 1.
DR   Pfam; PF00641; zf-RanBP; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS01358; ZF_RANBP2_1; 1.
DR   PROSITE; PS50199; ZF_RANBP2_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Nuclear protein; Ligase; Ubiquitin conjugation; Oncogene; Zinc;
KW   Zinc-finger; Metal-binding; Alternative splicing.
FT   DOMAIN       19    108       REGION I.
FT   DOMAIN      179    185       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      190    202       NUCLEAR EXPORT SEQUENCE.
FT   DOMAIN      210    304       ARF BINDING.
FT   DOMAIN      210    215       POLY-SER.
FT   DOMAIN      242    331       REGION II.
FT   DOMAIN      243    301       ASP/GLU-RICH (ACIDIC).
FT   ZN_FING     299    328       RANBP2-TYPE.
FT   ZN_FING     434    475       RING-TYPE.
FT   DOMAIN      462    469       NUCLEOLAR LOCALIZATION SIGNAL
FT                                (POTENTIAL).
FT   VARSPLIC      1     61       MISSING (IN ISOFORM MDM2-ALPHA).
FT   CONFLICT     11     11       G -> D (IN REF. 2).
FT   CONFLICT    238    239       QD -> HH (IN REF. 2).
SQ   SEQUENCE   487 AA;  54696 MW;  60CDB470A32A8E69 CRC64;
```

```
  Query Match             85.7%;  Score 30;  DB 1;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 57;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT   5
MDM2_MOUSE
ID   MDM2_MOUSE     STANDARD;      PRT;   489 AA.
AC   P23804; Q64330; Q61040;
DT   01-NOV-1991 (Rel. 20, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (P53-binding protein
DE   Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein).
GN   MDM2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM MDM2-P90).
RX   MEDLINE=91224107; PubMed=2026149;
RA   Fakharzadeh S.S., Trusko S.P., George D.L.;
RT   "Tumorigenic potential associated with enhanced expression of a gene
RT   that is amplified in a mouse tumor cell line.";
RL   EMBO J. 10:1565-1569(1991).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM MDM2-P90).
RC   STRAIN=129/SV;
RX   MEDLINE=97074674; PubMed=8917101;
RA   Jones S.N., Ansari-Lari M.A., Hancock A.R., Jones W.J., Gibbs R.A.,
RA   Donehower L.A., Bradley A.;
RT   "Genomic organization of the mouse double minute 2 gene.";
RL   Gene 175:209-213(1996).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM MDM2-P90).
RC   STRAIN=129/SV;
RX   MEDLINE=96299630; PubMed=8660994;
RA   de Oca Luna R.M., Tabor A.D., Eberspaecher H., Hulboy D.L.,
RA   Worth L.L., Colman M.S., Finlay C.A., Lozano G.;
RT   "The organization and expression of the mdm2 gene.";
RL   Genomics 33:352-357(1996).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS MDM2-P90 AND MDM2-P76).
RX   MEDLINE=99175199; PubMed=10075719;
RA   Saucedo L.J., Myers C.D., Perry M.E.;
RT   "Multiple murine double minute gene 2 (MDM2) proteins are induced by
RT   ultraviolet light.";
RL   J. Biol. Chem. 274:8161-8168(1999).
RN   [5]
RP   NUCLEOLAR LOCALIZATION SIGNAL.
RX   MEDLINE=20180080; PubMed=10713175;
RA   Weber J.D., Kuo M.-L., Bothner B., DiGiammarino E.L., Kriwacki R.W.,
RA   Roussel M.F., Sherr C.J.;
RT   "Cooperative signals governing ARF-mdm2 interaction and nucleolar
RT   localization of the complex.";
RL   Mol. Cell. Biol. 20:2517-2528(2000).
RN   [6]
RP   PHOSPHORYLATION BY ATM.
RX   MEDLINE=20079591; PubMed=10611322;
RA   Khosravi R., Maya R., Gottlieb T., Oren M., Shiloh Y., Shkedy D.;
RT   "Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation
RT   in response to DNA damage.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:14973-14977(1999).
CC   -!- FUNCTION: INHIBITS P53-AND P73-MEDIATED CELL CYCLE ARREST AND
CC       APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
CC       FUNCTIONS AS AN UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
```

```
CC       TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
CC       TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS.
CC   -!- COFACTOR: ZINC IS REQUIRED FOR UBIQUITIN LIGASE E3 ACTIVITY.
CC   -!- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
CC       SPECIFICALLY TO RNA. CAN INTERACTS ALSO WITH RETINOBLASTOMA
CC       PROTEIN (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1
CC       TRANSCRIPTION FACTOR.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC. EXPRESSED
CC       PREDOMINANTLY IN THE NUCLEOPLASM. INTERACTION WITH ARF(P14)
CC       RESULTS IN THE LOCALIZATION OF BOTH PROTEINS TO THE NUCLEOLUS. THE
CC       NUCLEOLAR LOCALIZATION SIGNALS IN BOTH ARF(P14) AND MDM2 MAY BE
CC       NECESSARY TO ALLOW EFFICIENT NUCLEOLAR LOCALIZATION OF BOTH
CC       PROTEINS.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; MDM2-P90 (SHOWN HERE) AND MDM2-
CC       P76; ARE PRODUCED BY ALTERNATIVE SPLICING AND ALSO BY ALTERNATIVE
CC       INITIATION.
CC   -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED AT LOW-LEVEL THROUGHOUT
CC       EMBRYO DEVELOPMENT AND IN ADULT TISSUES. MDM2-P90 IS MUCH MORE
CC       ABUNDANT THAN MDM2-P76 IN TESTIS, BRAIN, HEART, AND KIDNEY, BUT IN
CC       THE THYMUS, SPLEEN, AND INTESTINE, THE LEVELS OF THE MDM2 PROTEINS
CC       ARE ROUGHLY EQUIVALENT.
CC   -!- INDUCTION: BY UV LIGHT.
CC   -!- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
CC       G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
CC       REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
CC       INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
CC       FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
CC       ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
CC       PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
CC       ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
CC       ITSELF.
CC   -!- PTM: PHOSPHORYLATED IN RESPONSE TO IONIZING RADIATION IN AN ATM-
CC       DEPENDENT MANNER.
CC   -!- DISEASE: THE GENE FOR THIS PROTEIN IS AMPLIFIED IN A MOUSE TUMOR
CC       CELL LINE.
CC   -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC   -!- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
CC   -!- SIMILARITY: BELONGS TO THE MDM2 / MDM4 FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X58876; CAA41684.1; -.
DR   EMBL; U40145; AAA91167.1; -.
DR   EMBL; U47944; AAB09030.1; -.
DR   EMBL; U47935; AAB09030.1; JOINED.
DR   EMBL; U47936; AAB09030.1; JOINED.
DR   EMBL; U47937; AAB09030.1; JOINED.
DR   EMBL; U47938; AAB09030.1; JOINED.
DR   EMBL; U47939; AAB09030.1; JOINED.
DR   EMBL; U47940; AAB09030.1; JOINED.
DR   EMBL; U47941; AAB09030.1; JOINED.
DR   EMBL; U47942; AAB09030.1; JOINED.
DR   EMBL; U47943; AAB09030.1; JOINED.
DR   EMBL; U47934; AAB09031.1; -.
DR   PIR; S15349; S15349.
DR   HSSP; Q9UMT8; 1YCR.
DR   MGD; MGI:96952; Mdm2.
DR   InterPro; IPR003160; MDM2.
DR   InterPro; IPR001876; Znf-RanBP.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF02279; MDM2; 1.
DR   Pfam; PF00641; zf-RanBP; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS01358; ZF_RANBP2_1; 1.
DR   PROSITE; PS50199; ZF_RANBP2_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Nuclear protein; Ligase; Ubiquitin conjugation; Oncogene;
KW   Alternative splicing; Alternative initiation; Zinc; Zinc-finger;
KW   Metal-binding; Phosphorylation.
FT   MOD_RES       1      1       BLOCKED.
FT   DOMAIN       19    108       REGION I.
FT   DOMAIN      176    182       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      183    195       NUCLEAR EXPORT SEQUENCE.
FT   DOMAIN      203    213       POLY-SER.
FT   DOMAIN      208    302       ARF BINDING.
FT   DOMAIN      240    329       REGION II.
FT   DOMAIN      221    299       ASP/GLU-RICH (ACIDIC).
FT   ZN_FING     297    326       RANBP2-TYPE.
FT   ZN_FING     436    477       RING-TYPE.
FT   DOMAIN      464    471       NUCLEOLAR LOCALIZATION SIGNAL
FT                                (POTENTIAL).
FT   CONFLICT    203    203       S -> T (IN REF. 1).
FT   CONFLICT    419    419       D -> H (IN REF. 1).
FT   CONFLICT    486    486       S -> T (IN REF. 3).
SQ   SEQUENCE   489 AA;  54543 MW;  4ABF489E92038DF4 CRC64;


  Query Match             85.7%;  Score 30;  DB 1;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT   6
MDM2_HORSE
ID   MDM2_HORSE      STANDARD;      PRT;   491 AA.
AC   P56951;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (P53-binding protein
DE   Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (Edm2).
GN   MDM2.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20218866; PubMed=10754200;
RA   Nasir L., Burr P.D., McFarlane S.T., Gault E., Thompson H.,
RA   Argyle D.J.;
RT   "Cloning, sequence analysis and expression of the cDNAs encoding the
RT   canine and equine homologues of the mouse double minute 2 (mdm2)
RT   proto-oncogene.";
RL   Cancer Lett. 152:9-13(2000).
CC   -!- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
CC       APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
CC       FUNCTIONS AS AN UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
CC       TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
CC       TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS (BY SIMILARITY).
CC   -!- COFACTOR: ZINC IS REQUIRED FOR UBIQUITIN LIGASE E3 ACTIVITY (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
CC       SPECIFICALLY TO RNA. CAN INTERACTS ALSO WITH RETINOBLASTOMA
CC       PROTEIN (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1
CC       TRANSCRIPTION FACTOR (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC. EXPRESSED
CC       PREDOMINANTLY IN THE NUCLEOPLASM (BY SIMILARITY).
CC   -!- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
CC       G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
CC       REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
CC       INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
CC       FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
CC       ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
CC       PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
```

```
CC       ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
CC       ITSELF (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC   -!- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
CC   -!- SIMILARITY: BELONGS TO THE MDM2 / MDM4 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF121140; AAF28866.1; -.
DR   HSSP; Q9UMT8; 1YCR.
DR   InterPro; IPR003160; MDM2.
DR   InterPro; IPR001876; Znf-RanBP.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF02279; MDM2; 1.
DR   Pfam; PF00641; zf-RanBP; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS01358; ZF_RANBP2_1; 1.
DR   PROSITE; PS50199; ZF_RANBP2_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Nuclear protein; Ligase; Ubiquitin conjugation; Oncogene; Zinc;
KW   Zinc-finger; Metal-binding.
FT   DOMAIN       19    108       REGION I.
FT   DOMAIN      179    185       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      190    202       NUCLEAR EXPORT SEQUENCE.
FT   DOMAIN      210    304       ARF BINDING.
FT   DOMAIN      210    215       POLY-SER.
FT   DOMAIN      242    331       REGION II.
FT   DOMAIN      243    301       ASP/GLU-RICH (ACIDIC).
FT   ZN_FING     299    328       RANBP2-TYPE.
FT   ZN_FING     438    479       RING-TYPE.
FT   DOMAIN      466    473       NUCLEOLAR LOCALIZATION SIGNAL
FT                                (POTENTIAL).
SQ   SEQUENCE   491 AA;  55279 MW;  641E033D5C1DEC39 CRC64;


  Query Match             85.7%;  Score 30;  DB 1;  Length 491;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT   7
MDM2_HUMAN
ID   MDM2_HUMAN     STANDARD;      PRT;   491 AA.
AC   Q00987; Q13226; Q13297; Q13298; Q13299; Q13300; Q13301; Q9UGI3;
AC   Q9UMT8;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (P53-binding protein
DE   Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (Hdm2).
GN   MDM2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM MDM2).
RX   MEDLINE=92310576; PubMed=1614537;
RA   Oliner J.D., Kinzler K.W., Meltzer P.S., George D.L.,
RA   Vogelstein B.;
RT   "Amplification of a gene encoding a p53-associated protein in human
RT   sarcomas.";
RL   Nature 358:80-83(1992).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS MDM2-A; -B; -C; -D AND -E).
RC   TISSUE=Ovarian carcinoma;
RX   MEDLINE=96313107; PubMed=8705862;
RA   Sigalas I., Calvert A.H., Anderson J.J., Neal D.E., Lunec J.;
RT   "Alternatively spliced mdm2 transcripts with loss of p53 binding
RT   domain sequences: transforming ability and frequent detection in human
RT   cancer.";
RL   Nat. Med. 2:912-917(1996).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM MDM2-ALPHA).
RX   MEDLINE=20065171; PubMed=10597303;
RA   Veldhoen N., Metcalfe S., Milner J.;
RT   "A novel exon within the mdm2 gene modulates translation initiation in
RT   vitro and disrupts the p53-binding domain of mdm2 protein.";
RL   Oncogene 18:7026-7033(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM MDM2).
RC   TISSUE=Muscle;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 6-491 FROM N.A. (ISOFORM MDM2-A1).
RA   Liang H., Atkins H., Abdel-Fattah R., Suaeyun R., Lunec J.;
RT   "Genomic Organisation of the Human MDM2 Oncogene and Relationship to
RT   its Alternatively Spliced mRNA's.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 1-24 FROM N.A.
RX   MEDLINE=95380270; PubMed=7651818;
RA   Zauberman A., Flusberg D., Haupt Y., Barak Y., Oren M.;
RT   "A functional p53-responsive intronic promoter is contained within
RT   the human mdm2 gene.";
RL   Nucleic Acids Res. 23:2584-2592(1995).
RN   [7]
RP   SEQUENCE OF 1-9 FROM N.A.
RX   MEDLINE=97413643; PubMed=9270029;
RA   Landers J.E., Cassel S.L., George D.L.;
RT   "Translational enhancement of mdm2 oncogene expression in human tumor
RT   cells containing a stabilized wild-type p53 protein.";
RL   Cancer Res. 57:3562-3568(1997).
RN   [8]
RP   SEQUENCE OF 301-481 FROM N.A.
RX   MEDLINE=20542019; PubMed=11087894;
RA   Taubert H., Kappler M., Meye A., Bartel F., Schlott T.,
RA   Lautenschlaeger C., Bache M., Schmidt H., Wuerl P.;
RT   "A MboII polymorphism in exon 11 of the human MDM2 gene occuring in
RT   normal blood donors and in soft tissue sarcoma patients: an
RT   indication for an increased cancer susceptibility?";
RL   Mutat. Res. 456:39-44(2000).
RN   [9]
RP   MUTAGENESIS OF CYS-464.
RX   MEDLINE=98111004; PubMed=9450543;
RA   Honda R., Tanaka H., Yasuda H.;
RT   "Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53.";
RL   FEBS Lett. 420:25-27(1997).
RN   [10]
RP   MUTAGENESIS OF CYS-449.
RX   MEDLINE=20190101; PubMed=10723139;
RA   Honda R., Yasuda H.;
RT   "Activity of MDM2, a ubiquitin ligase, toward p53 or itself is
RT   dependent on the RING finger domain of the ligase.";
RL   Oncogene 19:1473-1476(2000).
RN   [11]
RP   MUTAGENESIS.
RX   MEDLINE=20187618; PubMed=10722742;
RA   Fang S., Jensen J.P., Ludwig R.L., Vousden K.H., Weissman A.M.;
RT   "Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself
RT   and p53.";
RL   J. Biol. Chem. 275:8945-8951(2000).
RN   [12]
RP   MUTAGENESIS OF CYS-441 AND CYS-478.
```

```
RX   MEDLINE=20076498; PubMed=10608892;
RA   Sharp D.A., Kratowicz S.A., Sank M.J., George D.L.;
RT   "Stabilization of the MDM2 oncoprotein by interaction with the
RT   structurally related MDMX protein.";
RL   J. Biol. Chem. 274:38189-38196(1999).
RN   [13]
RP   NUCLEOLAR LOCALIZATION SIGNAL.
RX   MEDLINE=20173879; PubMed=10707090;
RA   Lohrum M.A.E., Ashcroft M., Kubbutat M.H.G., Vousden K.H.;
RT   "Identification of a cryptic nucleolar-localization signal in MDM2.";
RL   Nat. Cell Biol. 2:179-181(2000).
RN   [14]
RP   PHOSPHORYLATION BY ATM.
RX   MEDLINE=20079591; PubMed=10611322;
RA   Khosravi R., Maya R., Gottlieb T., Oren M., Shiloh Y., Shkedy D.;
RT   "Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation
RT   in response to DNA damage.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:14973-14977(1999).
RN   [15]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 25-109 IN COMPLEX WITH P53.
RX   MEDLINE=97081050; PubMed=8875929;
RA   Kussie P.H., Gorina S., Marechal V., Elenbaas B., Moreau J.,
RA   Levine A.J., Pavletich N.P.;
RT   "Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor
RT   transactivation domain.";
RL   Science 274:948-953(1996).
CC   -!- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
CC       APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
CC       FUNCTIONS AS AN UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
CC       TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
CC       TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS.
CC   -!- COFACTOR: ZINC IS REQUIRED FOR UBIQUITIN LIGASE E3 ACTIVITY.
CC   -!- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
CC       SPECIFICALLY TO RNA. CAN INTERACTS ALSO WITH RETINOBLASTOMA
CC       PROTEIN (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1
CC       TRANSCRIPTION FACTOR.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC. EXPRESSED
CC       PREDOMINANTLY IN THE NUCLEOPLASM. INTERACTION WITH ARF(P14)
CC       RESULTS IN THE LOCALIZATION OF BOTH PROTEINS TO THE NUCLEOLUS. THE
CC       NUCLEOLAR LOCALIZATION SIGNALS IN BOTH ARF(P14) AND MDM2 MAY BE
CC       NECESSARY TO ALLOW EFFICIENT NUCLEOLAR LOCALIZATION OF BOTH
CC       PROTEINS.
CC   -!- ALTERNATIVE PRODUCTS: 8 ISOFORMS; MDM2 (SHOWN HERE), MDM2-A, MDM2-
CC       A1, MDM2-B, MDM2-C, MDM2-D, MDM2-E AND MDM2-ALPHA; ARE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: UBIQUITOUS. ISOFORMS MDM2-A, -B, -C, -D AND -E
CC       ARE OBSERVED IN A RANGE OF HUMAN CANCERS BUT ABSENT IN NORMAL
CC       TISSUES.
CC   -!- INDUCTION: BY DNA DAMAGE.
CC   -!- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
CC       G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
CC       REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
CC       INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
CC       FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
CC       ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
CC       PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
CC       ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
CC       ITSELF.
CC   -!- PTM: PHOSPHORYLATED IN RESPONSE TO IONIZING RADIATION IN AN ATM-
CC       DEPENDENT MANNER.
CC   -!- DISEASE: SEEMS TO BE AMPLIFIED IN CERTAIN TUMORS (INCLUDING SOFT
CC       TISSUE SARCOMAS, OSTEOSARCOMAS AND GLIOMAS). A HIGHER FREQUENCY OF
CC       SPLICE VARIANTS LACKING P53 BINDING DOMAIN SEQUENCES WAS FOUND IN
CC       LATE-STAGE AND HIGH-GRADE OVARIAN AND BLADDER CARCINOMAS. FOUR OF
CC       THE SPLICE VARIANTS SHOW LOSS OF P53 BINDING.
CC   -!- MISCELLANEOUS: MDM2 RING FINGER MUTATIONS THAT FAILED TO
CC       UBIQUITINATE P53 IN VITRO DID NOT TARGET P53 FOR DEGRADATION WHEN
CC       EXPRESSED IN CELLS.
CC   -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC   -!- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
CC   -!- SIMILARITY: BELONGS TO THE MDM2 / MDM4 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M92424; AAA60568.1; -.
DR   EMBL; Z12020; CAA78055.1; -.
DR   EMBL; U33199; AAA75514.1; -.
DR   EMBL; U33200; AAA75515.1; -.
DR   EMBL; U33201; AAA75516.1; -.
DR   EMBL; U33202; AAA75517.1; -.
DR   EMBL; U33203; AAA75518.1; -.
DR   EMBL; BC009893; AAH09893.1; -.
DR   EMBL; U28935; AAA82237.1; -.
DR   EMBL; U39736; AAA82061.1; -.
DR   EMBL; AJ251943; CAB64448.1; -.
DR   PIR; S24354; S24354.
DR   PDB; 1YCR; 19-NOV-97.
DR   MIM; 164785; -.
DR   InterPro; IPR003160; MDM2.
DR   InterPro; IPR001876; Znf-RanBP.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF02279; MDM2; 1.
DR   Pfam; PF00641; zf-RanBP; 1.
DR   PROSITE; PS01358; ZF_RANBP2_1; 1.
DR   PROSITE; PS50199; ZF_RANBP2_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Nuclear protein; Ligase; Ubiquitin conjugation; Oncogene;
KW   Alternative splicing; Zinc; Zinc-finger; Metal-binding;
KW   Phosphorylation; 3D-structure.
FT   DOMAIN       19    108       REGION I.

  Query Match             85.7%;  Score 30;  DB 1;  Length 491;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT   8
YD56_SCHPO
ID   YD56_SCHPO     STANDARD;      PRT;   515 AA.
AC   Q10310;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Hypothetical 59.0 kDa protein C6C3.06 in chromosome I.
GN   SPAC6C3.06.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z69731; CAA93619.1; -.
```

```
KW   Hypothetical protein.
SQ   SEQUENCE   515 AA;  59035 MW;  B0466464840B3F06 CRC64;


  Query Match             85.7%;  Score 30;  DB 1;  Length 515;
  Best Local Similarity   87.5%;  Pred. No. 61;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          |||||| |
Db     49 KLLLKLTK 56


RESULT   9
GGA3_HUMAN
ID   GGA3_HUMAN     STANDARD;      PRT;   723 AA.
AC   Q9NZ52; Q9UJY3; Q15017;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   ADP-ribosylation factor binding protein GGA3 (Golgi-localized, gamma
DE   ear-containing, ARF-binding protein 3).
GN   GGA3 OR KIAA0154.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RC   TISSUE=Heart;
RX   MEDLINE=20211638; PubMed=10747089;
RA   Dell'Angelica E.C., Puertollano R., Mullins C., Aguilar R.C.,
RA   Vargas J.D., Hartnell L.M., Bonifacino J.S.;
RT   "GGAs: a family of ADP ribosylation factor-binding proteins related to
RT   adaptors and associated with the Golgi complex.";
RL   J. Cell Biol. 149:81-94(2000).
RN   [2]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RX   MEDLINE=20214818; PubMed=10749927;
RA   Boman A.L., Zhang C.-J., Zhu X., Kahn R.A.;
RT   "A family of ADP-ribosylation factor effectors that can alter
RT   transport through the trans-Golgi.";
RL   Mol. Biol. Cell 11:1241-1255(2000).
RN   [3]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   TISSUE=Bone marrow;
RX   MEDLINE=96127530; PubMed=8590280;
RA   Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RT   "Prediction of the coding sequences of unidentified human genes. IV.
RT   The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT   analysis of cDNA clones from human cell line KG-1.";
RL   DNA Res. 2:167-174(1995).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF MEMBRANE TRAFFIC
CC       THROUGH THE TRANS-GOLGI NETWORK.
CC   -!- SUBUNIT: BINDS TO ARF1.
CC   -!- SUBCELLULAR LOCATION: TRANS-GOLGI NETWORK.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.
CC   -!- SIMILARITY: CONTAINS 1 GAMMA-ADAPTIN C-TERMINAL DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 VHS DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF219138; AAF42848.1; -.
DR   EMBL; AF190864; AAF05709.1; -.
DR   EMBL; AF219139; AAF42849.1; -.
DR   EMBL; D63876; BAA09926.1; ALT_INIT.
DR   InterPro; IPR001121; G_adapt_C.
DR   InterPro; IPR002014; HRS.
DR   Pfam; PF00790; VHS; 1.
DR   ProDom; PD003686; HRS; 1.
DR   ProDom; PD021457; G_adapt_C; 1.
DR   SMART; SM00288; VHS; 1.
DR   PROSITE; PS50179; VHS; 1.
KW   Protein transport; Golgi stack; Coiled coil; Alternative splicing.
FT   DOMAIN        1    313       BINDS TO ARF1 (IN LONG ISOFORM).
FT   DOMAIN       16    146       VHS.
FT   DOMAIN      188    232       COILED COIL (POTENTIAL).
FT   DOMAIN      357    360       POLY-PRO.
FT   DOMAIN      453    457       POLY-SER.
FT   DOMAIN      624    629       POLY-VAL.
FT   DOMAIN      598    709       GAMMA-ADAPTIN C-TERMINAL.
FT   VARSPLIC     68    100       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   723 AA;  78314 MW;  4F80D6032239168C CRC64;


  Query Match             85.7%;  Score 30;  DB 1;  Length 723;
  Best Local Similarity   87.5%;  Pred. No. 85;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          ||| ||||
Db    176 KLLAKLLK 183


RESULT  10
DPOL_NPVAC
ID   DPOL_NPVAC     STANDARD;      PRT;   984 AA.
AC   P18131;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   DNA polymerase (EC 2.7.7.7).
GN   POL.
OS   Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=46015;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L1;
RX   MEDLINE=89073763; PubMed=3059678;
RA   Tomalski M.D., Wu J.G., Miller L.K.;
RT   "The location, sequence, transcription, and regulation of a
RT   baculovirus DNA polymerase gene.";
RL   Virology 167:591-600(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C6;
RX   MEDLINE=94303173; PubMed=8030224;
RA   Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT   "The complete DNA sequence of Autographa californica nuclear
RT   polyhedrosis virus.";
RL   Virology 202:586-605(1994).
CC   -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC       + {DNA}(N).
CC   -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M20744; AAA46692.1; -.
DR   EMBL; L22858; AAA66695.1; -.
DR   PIR; A31832; DJNVCP.
```

```
DR   InterPro; IPR002064; DNA_pol_B.
DR   Pfam; PF00136; DNA_pol_B; 1.
DR   Pfam; PF03104; DNA_pol_B_exo; 1.
DR   PRINTS; PR00106; DNAPOLB.
DR   SMART; SM00486; POLBc; 1.
DR   PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW   Transferase; DNA-directed DNA polymerase; DNA replication;
KW   DNA-binding; Early protein.
FT   DOMAIN      724    727       POLY-LYS.
FT   DOMAIN      946    960       POLY-ASP.
FT   CONFLICT    830    830       R -> W (IN REF. 1).
SQ   SEQUENCE   984 AA;  114307 MW;  156ABB6BA1B45A21 CRC64;


  Query Match             85.7%;  Score 30;  DB 1;  Length 984;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLKLL 7
          |||||||
Db    569 KLLLKLL 575


RESULT  11
DPOL_NPVBM
ID   DPOL_NPVBM     STANDARD;      PRT;   986 AA.
AC   P41712; O92430;
DT   01-NOV-1995 (Rel. 32, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   DNA polymerase (EC 2.7.7.7).
GN   POL.
OS   Bombyx mori nuclear polyhedrosis virus (BmNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=10458;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95133178; PubMed=7831799;
RA   Chaeychomsri S., Ikeda M., Kobayashi M.;
RT   "Nucleotide sequence and transcriptional analysis of the DNA
RT   polymerase gene of Bombyx mori nuclear polyhedrosis virus.";
RL   Virology 206:435-447(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=T3;
RA   Gomi S., Majima K., Maeda S.;
RT   "Sequence analysis of the genome of Bombyx mori
RT   nucleopolyhedrovirus.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC       + {DNA}(N).
CC   -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D16231; BAA03756.1; -.
DR   EMBL; L33180; AAC63738.1; -.
DR   InterPro; IPR002064; DNA_pol_B.
DR   Pfam; PF00136; DNA_pol_B; 1.
DR   Pfam; PF03104; DNA_pol_B_exo; 1.
DR   PRINTS; PR00106; DNAPOLB.
DR   SMART; SM00486; POLBc; 1.
DR   PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW   Transferase; DNA-directed DNA polymerase; DNA replication;
KW   DNA-binding; Early protein.
FT   DOMAIN      724    727       POLY-LYS.
FT   DOMAIN      947    951       POLY-ASP.
FT   DOMAIN      954    959       POLY-ASP.
FT   CONFLICT    116    116       A -> S (IN REF. 1).
FT   CONFLICT    245    245       H -> Y (IN REF. 1).
FT   CONFLICT    250    250       H -> Y (IN REF. 1).
FT   CONFLICT    258    258       V -> I (IN REF. 1).
FT   CONFLICT    478    479       TA -> AG (IN REF. 1).
FT   CONFLICT    941    941       S -> G (IN REF. 1).
FT   CONFLICT    952    952       N -> NDN (IN REF. 1).
SQ   SEQUENCE   986 AA;  114418 MW;  503E39F4A0BCC125 CRC64;


  Query Match             85.7%;  Score 30;  DB 1;  Length 986;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLKLL 7
          |||||||
Db    569 KLLLKLL 575


RESULT  12
CAPG_STAAU
ID   CAPG_STAAU     STANDARD;      PRT;   172 AA.
AC   P39856;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Protein capG.
GN   CAPG.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M;
RX   MEDLINE=95050273; PubMed=7961465;
RA   Lin W.S., Cunneen T., Lee C.Y.;
RT   "Sequence analysis and molecular characterization of genes required
RT   for the biosynthesis of type 1 capsular polysaccharide in
RT   Staphylococcus aureus.";
RL   J. Bacteriol. 176:7005-7016(1994).
CC   -!- FUNCTION: REQUIRED FOR THE BIOSYNTHESIS OF TYPE 1 CAPSULAR
CC       POLYSACCHARIDE.
CC   -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC       ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U10927; AAA64646.1; -.
DR   InterPro; IPR001451; Hexapep_transf.
DR   Pfam; PF00132; hexapep; 3.
DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW   Transferase; Repeat.
SQ   SEQUENCE   172 AA;  19451 MW;  1608180D13A10EA4 CRC64;


  Query Match             82.9%;  Score 29;  DB 1;  Length 172;
  Best Local Similarity   75.0%;  Pred. No. 33;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          || |||:|
Db      3 KLFLKLMK 10
```

```
RESULT  13
SOML_TETMU
ID   SOML_TETMU     STANDARD;      PRT;   229 AA.
AC   Q9I9H4;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Somatolactin precursor (SL).
OS   Tetraodon miurus (Congo puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=94908;
RN   [1].
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RA   Rand-Weaver M., May D.;
RT   "Cloning and sequencing of Tetraodon miurus somatolactin.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF253066; AAF64522.1; -.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    229       SOMATOLACTIN.
FT   DISULFID     26     36       BY SIMILARITY.
FT   DISULFID     87    203       BY SIMILARITY.
FT   DISULFID    220    228       BY SIMILARITY.
FT   CARBOHYD    143    143       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   229 AA;  26125 MW;  C10CCF295D28C447 CRC64;


  Query Match           82.9%;  Score 29;  DB 1;  Length 229;
  Best Local Similarity   75.0%;  Pred. No. 44;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            ::||||||
Db     212 EILLKLLK 219


RESULT  14
SOM1_SPAAU
ID   SOM1_SPAAU     STANDARD;      PRT;   231 AA.
AC   P54863;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Somatolactin 1 precursor (SL).
OS   Sparus aurata (Gilthead sea bream).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sparidae; Sparus.
OX   NCBI_TaxID=8175;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=97114243; PubMed=8954766;
RA   Astola A., Pendon C., Ortiz M., Valdivia M.M.;
RT   "Cloning and expression of somatolactin, a pituitary hormone related
RT   to growth hormone and prolactin from gilthead seabream, Sparus
RT   aurata.";
RL   Gen. Comp. Endocrinol. 104:330-336(1996).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L49205; AAA98734.1; -.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    231       SOMATOLACTIN 1.
FT   DISULFID     29     39       BY SIMILARITY.
FT   DISULFID     89    205       BY SIMILARITY.
FT   DISULFID    222    230       BY SIMILARITY.
FT   CARBOHYD    145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   231 AA;  26961 MW;  67A44E7D43E02504 CRC64;


  Query Match           82.9%;  Score 29;  DB 1;  Length 231;
  Best Local Similarity   75.0%;  Pred. No. 45;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            ::||||||
Db     214 EILLKLLK 221


RESULT  15
SOM2_SPAAU
ID   SOM2_SPAAU     STANDARD;      PRT;   231 AA.
AC   P79894;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Somatolactin 2 precursor (SL).
OS   Sparus aurata (Gilthead sea bream).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sparidae; Sparus.
OX   NCBI_TaxID=8175;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RA   Cavari B., Funkenstein B., Kawauchi H.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------------
DR   EMBL; Y11144; CAA72031.1; -.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    231       SOMATOLACTIN 2.
FT   DISULFID     29     39       BY SIMILARITY.
FT   DISULFID     89    205       BY SIMILARITY.
FT   DISULFID    222    230       BY SIMILARITY.
FT   CARBOHYD    145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   231 AA;  26765 MW;  09C774C68DE08BA1 CRC64;


  Query Match             82.9%;  Score 29;  DB 1;  Length 231;
  Best Local Similarity   75.0%;  Pred. No. 45;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            ::||||||
Db     214 EILLKLLK 221


Search completed: June 17, 2002, 12:44:47
Job time: 302 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:44:20 ; Search time 73.61 Seconds
(without alignments)
18.801 Million cell updates/sec

Title: US-09-367-714A-29
Perfect score: 35
Sequence: 1 KLLLKLLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	32	91.4	302	5	Q26784	Q26784 trypanosoma
2	32	91.4	542	12	Q9IZC1	Q9izc1 salem virus
3	31	88.6	194	11	Q91YF2	Q91yf2 mus musculu
4	31	88.6	236	10	Q948F3	Q948f3 oryza sativ
5	31	88.6	296	16	Q92HH2	Q92hh2 rickettsia
6	31	88.6	359	11	O88829	O88829 mus musculu
7	31	88.6	359	11	Q9QWF8	Q9qwf8 mus musculu
8	31	88.6	384	17	Q9HK94	Q9hk94 thermoplasm
9	31	88.6	387	11	Q9CZ65	Q9cz65 mus musculu
10	31	88.6	387	11	Q9QWF9	Q9qwf9 mus musculu
11	31	88.6	960	12	Q91GJ3	Q91gj3 epiphyas po
12	31	88.6	1211	13	Q91086	Q91086 meleagris g
13	30	85.7	60	4	Q96DS5	Q96ds5 homo sapien
14	30	85.7	66	4	Q96DS3	Q96ds3 homo sapien
15	30	85.7	95	4	Q96DS1	Q96ds1 homo sapien
16	30	85.7	110	2	Q9R2T1	Q9r2t1 borrelia bu
17	30	85.7	110	2	Q9S089	Q9s089 borrelia bu
18	30	85.7	110	2	Q9S047	Q9s047 borrelia bu
19	30	85.7	110	2	Q9RZZ6	Q9rzz6 borrelia bu
20	30	85.7	110	2	Q44792	Q44792 borrelia bu
21	30	85.7	130	4	Q9H4C3	Q9h4c3 homo sapien
22	30	85.7	135	5	Q95WZ4	Q95wz4 ixodes scap
23	30	85.7	137	11	Q9DBL5	Q9dbl5 mus musculu
24	30	85.7	159	4	Q96DS0	Q96ds0 homo sapien
25	30	85.7	162	8	Q950Q3	Q950q3 spizellomyc
26	30	85.7	166	16	Q9PPQ9	Q9ppq9 ureaplasma
27	30	85.7	191	17	Q97WI9	Q97wi9 sulfolobus
28	30	85.7	195	4	Q96DS4	Q96ds4 homo sapien
29	30	85.7	217	10	Q942V5	Q942v5 oryza sativ
30	30	85.7	237	4	Q96LK7	Q96lk7 homo sapien
31	30	85.7	243	4	Q9H4C5	Q9h4c5 homo sapien
32	30	85.7	282	16	Q9ZLZ4	Q9zlz4 helicobacte
33	30	85.7	321	17	O58243	O58243 pyrococcus
34	30	85.7	325	13	Q9PVL2	Q9pvl2 gallus gall
35	30	85.7	396	4	Q9NXT4	Q9nxt4 homo sapien
36	30	85.7	418	10	Q9LDC5	Q9ldc5 oryza sativ
37	30	85.7	487	6	Q9GMZ6	Q9gmz6 canis famil
38	30	85.7	487	6	Q95KN5	Q95kn5 canis famil
39	30	85.7	489	11	Q91XK7	Q91xk7 mus musculu
40	30	85.7	586	4	Q9NRU3	Q9nru3 homo sapien
41	30	85.7	586	11	Q9JIQ6	Q9jiq6 mus musculu
42	30	85.7	631	4	Q9NRN1	Q9nrn1 homo sapien
43	30	85.7	633	4	Q9NRK5	Q9nrk5 homo sapien
44	30	85.7	633	4	Q9H952	Q9h952 homo sapien
45	30	85.7	644	11	Q9JIM7	Q9jim7 mus musculu

ALIGNMENTS

RESULT 1
Q26784
ID Q26784 PRELIMINARY; PRT; 302 AA.
AC Q26784;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MRNA SEQUENCE (FRAGMENT).
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95140053; PubMed=7838181;
RA Woodward R., Carden M.J., Gull K.;
RT "Molecular characterisation of a novel, repetitive protein of the
RT paraflagellar rod in Trypanosoma brucei.";
RL Mol. Biochem. Parasitol. 67:31-39(1994).
DR EMBL; M87318; AAC37211.1; -.
FT NON_TER 1 1
SQ SEQUENCE 302 AA; 36223 MW; B3CA5AB868E877CC CRC64;

Query Match 91.4%; Score 32; DB 5; Length 302;
Best Local Similarity 87.5%; Pred. No. 98;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLLKLLK 8
||||||:||
Db 21 KLLLKVLK 28

RESULT 2
Q9IZC1
ID Q9IZC1 PRELIMINARY; PRT; 542 AA.
AC Q9IZC1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

```
DE   N PROTEIN.
GN   N.
OS   Salem virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Paramyxovirus.
OX   NCBI_TaxID=120499;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20255542; PubMed=10793001;
RA   Renshaw R.W., Glaser A.L., Van Campen H., Weiland F., Dubovi E.J.;
RT   "Identification and phylogenetic comparison of salem virus, a novel
RT   paramyxovirus of horses.";
RL   Virology 270:417-429(2000).
DR   EMBL; AF237881; AAF63741.1; -.
DR   InterPro; IPR002021; Paramyx_ncap.
DR   Pfam; PF00973; Paramyx_ncap; 1.
SQ   SEQUENCE   542 AA;  60717 MW;  F057ECF483F1F7D9 CRC64;


  Query Match             91.4%;  Score 32;  DB 12;  Length 542;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          |||||||:
Db     50 KLLLKLLR 57


RESULT   3
Q91YF2
ID   Q91YF2      PRELIMINARY;      PRT;   194 AA.
AC   Q91YF2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GM3 SYNTHASE PROTEIN.
GN   GM3 SYNTHASE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR; TISSUE=BRAIN;
RA   Shuichi T.;
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y18023; CAC79655.1; -.
SQ   SEQUENCE   194 AA;  22205 MW;  E68780E30BFE84EC CRC64;


  Query Match             88.6%;  Score 31;  DB 11;  Length 194;
  Best Local Similarity   87.5%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          | ||||||
Db    175 KFLLKLLK 182


RESULT   4
Q948F3
ID   Q948F3      PRELIMINARY;      PRT;   236 AA.
AC   Q948F3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE RIBONUCLEOTIDE REDUCTASE (FRAGMENT).
GN   OSJNBA0049O12.22.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIPPONBARE;
RA   Spiegel L., Nascimento L., de la Bastide M., Kirchoff K., Preston R.,
RA   King L., Vil M.D., Baker J., Zutavern T., Santos L., Miller B.,
RA   Kuit K., Cunnius D.M., Balija V., Shah R., Bahret A., Bell M.,
RA   Yang C., Palmer L., O'Shaughnessy A., Dedhia N., McCombie W.R.;
RT   "Genomic Sequence for Oryza sativa, Nipponbare strain, clone
RT   OSJNBa0049O12, from chromosome 2, complete sequence.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC069158; AAK98710.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   236 AA;  26749 MW;  D29916BA145DC49A CRC64;


  Query Match             88.6%;  Score 31;  DB 10;  Length 236;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          |||||||:
Db      1 KLLLKLLQ 8


RESULT   5
Q92HH2
ID   Q92HH2      PRELIMINARY;      PRT;   296 AA.
AC   Q92HH2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL PROTEIN RC0799.
GN   RC0799.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MALISH 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
DR   EMBL; AE008635; AAL03337.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   296 AA;  33899 MW;  381A649E1DDBAB5C CRC64;


  Query Match             88.6%;  Score 31;  DB 16;  Length 296;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          |:|||:||
Db    287 KILLKILK 294


RESULT   6
O88829
ID   O88829      PRELIMINARY;      PRT;   359 AA.
AC   O88829;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   GM3 SYNTHASE (EC 2.4.99.9).
GN   SIAT9.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL6J; TISSUE=BRAIN;
RA   Ishii A., Saito M.;
RT   "Mouse GM3 Synthase cDNA.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB-C;
RA   Fukumoto S., Miyazaki H., Urano T., Furukawa K.;
RT   "Expression cloning of mouse cDNA of CMP-NeuAc: lactosylceramide
RT   alpha2,3sialyltransferase (GM3 synthase), the enzyme that initiates
RT   the synthesis of gangliosides).";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Kapitonov D., Yu R.K.;
RT   "Combinatorial PCR in homologous cloning: cloning of GM3 synthase (ST-
RT   I).";
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB018048; BAA33491.1; -.
DR   EMBL; AB013302; BAA76467.1; -.
DR   EMBL; AF119416; AAF66147.1; -.
DR   MGD; MGI:1339963; Siat9.
DR   InterPro; IPR001675; Glyco_transf_29.
DR   Pfam; PF00777; Glyco_transf_29; 1.
KW   Transferase; Glycosyltransferase.
SQ   SEQUENCE   359 AA;  41245 MW;  38D81D0B8CFC4961 CRC64;


  Query Match             88.6%;  Score 31;  DB 11;  Length 359;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
           | ||||||
Db     340 KFLLKLLK 347


RESULT   7
Q9QWF8
ID   Q9QWF8      PRELIMINARY;      PRT;   359 AA.
AC   Q9QWF8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   LACTOSYLCERAMIDE ALPHA-2,3-SIALYLTRANSFERASE (EC 2.4.99.9).
GN   SIAT9 OR ST3GAL V.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR; TISSUE=BRAIN;
RX   MEDLINE=99092398; PubMed=9875239;
RA   Kono M., Takashima S., Liu H., Inoue M., Kojima N., Young-Choon L.,
RA   Hamamoto T., Tsuji S.;
RT   "Molecular cloning and characterization of fifth type of beta-
RT   galactoside alpha-2,3-sialyltransferase (ST3Gal V; GM3 synthase).";
RL   Biochem. Biophys. Res. Commun. 253:170-175(1998).
DR   EMBL; Y15003; CAA75236.1; -.
DR   MGD; MGI:1339963; Siat9.
DR   InterPro; IPR001675; Glyco_transf_29.
DR   Pfam; PF00777; Glyco_transf_29; 1.
KW   Transferase; Glycosyltransferase.
SQ   SEQUENCE   359 AA;  41235 MW;  8E3C734CD1899E3C CRC64;


  Query Match             88.6%;  Score 31;  DB 11;  Length 359;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
           | ||||||
Db     340 KFLLKLLK 347


RESULT   8
Q9HK94
ID   Q9HK94      PRELIMINARY;      PRT;   384 AA.
AC   Q9HK94;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL PROTEIN TA0707.
GN   TA0707.
OS   Thermoplasma acidophilum.
OC   Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaceae;
OC   Thermoplasma.
OX   NCBI_TaxID=2303;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 1728;
RX   MEDLINE=20479972; PubMed=11029001;
RA   Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA   Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT   "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT   acidophilum.";
RL   Nature 407:508-513(2000).
DR   EMBL; AL445065; CAC11845.1; -.
DR   InterPro; IPR001296; Glycos_transf_1.
DR   Pfam; PF00534; Glycos_transf_1; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   384 AA;  44734 MW;  BEC8D05CF0237E40 CRC64;


  Query Match             88.6%;  Score 31;  DB 17;  Length 384;
  Best Local Similarity   75.0%;  Pred. No. 1.9e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
           |||:|:||
Db     218 KLLMKILK 225


RESULT   9
Q9CZ65
ID   Q9CZ65      PRELIMINARY;      PRT;   387 AA.
AC   Q9CZ65;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SIALYLTRANSFERASE 9 (CMP-NEUAC:LACTOSYLCERAMIDE ALPHA-2,3-
DE   SIALYLTRANSFERASE).
GN   SIAT9.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
```

```
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK012961; BAB28571.1; -.
DR   MGD; MGI:1339963; Siat9.
DR   InterPro; IPR001675; Glyco_transf_29.
DR   Pfam; PF00777; Glyco_transf_29; 1.
SQ   SEQUENCE   387 AA;  44572 MW;  7D358298034CDD96 CRC64;


  Query Match             88.6%;  Score 31;  DB 11;  Length 387;
  Best Local Similarity   87.5%;  Pred. No. 1.9e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          | ||||||
Db    368 KFLLKLLK 375


RESULT  10
Q9QWF9
ID   Q9QWF9      PRELIMINARY;      PRT;   387 AA.
AC   Q9QWF9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   LACTOSYLCERAMIDE ALPHA-2,3-SIALYLTRANSFERASE (EC 2.4.99.9).
GN   SIAT9 OR ST3GAL V.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR; TISSUE=BRAIN;
RX   MEDLINE=99092398; PubMed=9875239;
RA   Kono M., Takashima S., Liu H., Inoue M., Kojima N., Young-Choon L.,
RA   Hamamoto T., Tsuji S.;
RT   "Molecular cloning and characterization of fifth type of beta-
RT   galactoside alpha-2,3-sialyltransferase (ST3Gal V; GM3 synthase).";
RL   Biochem. Biophys. Res. Commun. 253:170-175(1998).
DR   EMBL; Y15003; CAA75235.1; -.
DR   MGD; MGI:1339963; Siat9.
DR   InterPro; IPR001675; Glyco_transf_29.
DR   Pfam; PF00777; Glyco_transf_29; 1.
KW   Transferase; Glycosyltransferase.
SQ   SEQUENCE   387 AA;  44562 MW;  CBD1ECDF5E390ACB CRC64;


  Query Match             88.6%;  Score 31;  DB 11;  Length 387;
  Best Local Similarity   87.5%;  Pred. No. 1.9e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          | ||||||
Db    368 KFLLKLLK 375


RESULT  11
Q91GJ3
ID   Q91GJ3      PRELIMINARY;      PRT;   960 AA.
AC   Q91GJ3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DNAPOL.
GN   DNAPOL.
OS   Epiphyas postvittana nucleopolyhedrovirus.
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=70600;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hyink O., Dellow R.A., Olsen M., Caradoc-Davies K.M.B., Drake K.,
RA   Ward V.K.;
RT   "The complete sequence of the Epiphyas postvittana
RT   nucleopolyhedrovirus genome.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY043265; AAK85622.1; -.
SQ   SEQUENCE   960 AA;  111520 MW;  0CB8A8E2E0F5B540 CRC64;


  Query Match             88.6%;  Score 31;  DB 12;  Length 960;
  Best Local Similarity   87.5%;  Pred. No. 4.3e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          ||||:|||
Db    562 KLLLELLK 569


RESULT  12
Q91086
ID   Q91086      PRELIMINARY;      PRT;  1211 AA.
AC   Q91086;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PHOSPHOLIPASE C BETA.
OS   Meleagris gallopavo (Common turkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX   NCBI_TaxID=9103;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96257751; PubMed=8687401;
RA   Waldo G.L., Paterson A., Boyer J.L., Nicholas R.A., Harden T.K.;
RT   "Molecular cloning, expression and regulatory activity of G alpha 11-
RT   and beta gamma-subunit-stimulated phospholipase C-beta from avian
RT   erythrocytes.";
RL   Biochem. J. 316:559-568(1996).
DR   EMBL; U49431; AAC60011.1; -.
DR   HSSP; P10688; 1QAS.
DR   InterPro; IPR000008; C2.
DR   InterPro; IPR001192; PI_PLC.
DR   InterPro; IPR000909; PI_PLC_X.
DR   InterPro; IPR001711; PI_PLC_Y.
DR   Pfam; PF00168; C2; 1.
DR   Pfam; PF00388; PI-PLC-X; 1.
DR   Pfam; PF00387; PI-PLC-Y; 1.
DR   PRINTS; PR00390; PHPHLIPASEC.
DR   ProDom; PD001202; PI_PLC_Y; 1.
DR   SMART; SM00239; C2; 1.
DR   SMART; SM00148; PLCXc; 1.
DR   SMART; SM00149; PLCYc; 1.
DR   PROSITE; PS50004; C2_DOMAIN_2; 1.
DR   PROSITE; PS50007; PIPLC_X_DOMAIN; 1.
DR   PROSITE; PS50008; PIPLC_Y_DOMAIN; 1.
SQ   SEQUENCE   1211 AA;  139061 MW;  4E96A10C6AFD6B5A CRC64;


  Query Match             88.6%;  Score 31;  DB 13;  Length 1211;
  Best Local Similarity   87.5%;  Pred. No. 5.2e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 KLLLKLLK 8
          || |||||
Db    902 KLFLKLLK 909


RESULT  13
```

```
Q96DS5
ID   Q96DS5      PRELIMINARY;      PRT;    60 AA.
AC   Q96DS5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MDM2 VARIANT FB25.
GN   MDM2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=RHABDOMYOSARCOMA TUMOR;
RA   Bartel F., Taylor A.C., Taubert H., Harris L.C.;
RT   "Novel mdm2 splice variants identified in pediatric rhabdomyosarcoma
RT   tumors and cell lines.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF385322; AAL13242.1; -.
SQ   SEQUENCE   60 AA;  6652 MW;  BCC2CC61C4CC98A3 CRC64;


  Query Match             85.7%;  Score 30;  DB 4;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 59;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT  14
Q96DS3
ID   Q96DS3      PRELIMINARY;      PRT;    66 AA.
AC   Q96DS3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MDM2 VARIANT FB28.
GN   MDM2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=RHABDOMYOSARCOMA TUMOR;
RA   Bartel F., Taylor A.C., Taubert H., Harris L.C.;
RT   "Novel mdm2 splice variants identified in pediatric rhabdomyosarcoma
RT   tumors and cell lines.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF385324; AAL13244.1; -.
SQ   SEQUENCE   66 AA;  7396 MW;  E3B3F3C385D4A8A5 CRC64;


  Query Match             85.7%;  Score 30;  DB 4;  Length 66;
  Best Local Similarity   100.0%;  Pred. No. 65;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT  15
Q96DS1
ID   Q96DS1      PRELIMINARY;      PRT;    95 AA.
AC   Q96DS1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MDM2 VARIANT FB30.
GN   MDM2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=RHABDOMYOSARCOMA TUMOR;
RA   Bartel F., Taylor A.C., Taubert H., Harris L.C.;
RT   "Novel mdm2 splice variants identified in pediatric rhabdomyosarcoma
RT   tumors and cell lines.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF385326; AAL13246.1; -.
SQ   SEQUENCE   95 AA;  10622 MW;  00052F95211E3612 CRC64;


  Query Match             85.7%;  Score 30;  DB 4;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


Search completed: June 17, 2002, 12:44:21
Job time: 296 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:42:04 ; Search time 34.71 Seconds
(without alignments)
5.630 Million cell updates/sec

Title: US-09-367-714A-29
Perfect score: 35
Sequence: 1 KLLLKLLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	88.6	414	4	US-09-334-601-4	Sequence 4, Appli
2	30	85.7	18	1	US-07-725-331-9	Sequence 9, Appli
3	30	85.7	18	5	PCT-US91-05047-9	Sequence 9, Appli
4	30	85.7	21	1	US-08-944-133-13	Sequence 13, Appl
5	30	85.7	86	2	US-08-248-839C-47	Sequence 47, Appl
6	30	85.7	489	1	US-07-903-103-4	Sequence 4, Appli
7	30	85.7	489	1	US-08-044-619A-4	Sequence 4, Appli
8	30	85.7	489	1	US-08-283-911-4	Sequence 4, Appli
9	30	85.7	489	1	US-08-245-500A-5	Sequence 5, Appli
10	30	85.7	489	1	US-08-390-546-5	Sequence 5, Appli
11	30	85.7	489	1	US-08-390-479A-5	Sequence 5, Appli
12	30	85.7	489	1	US-08-557-393-5	Sequence 5, Appli
13	30	85.7	489	1	US-08-390-516C-5	Sequence 5, Appli
14	30	85.7	489	1	US-08-390-517A-5	Sequence 5, Appli
15	30	85.7	489	1	US-08-390-515A-5	Sequence 5, Appli
16	30	85.7	489	2	US-08-801-718-5	Sequence 5, Appli
17	30	85.7	491	1	US-07-903-103-2	Sequence 2, Appli
18	30	85.7	491	1	US-08-044-619A-2	Sequence 2, Appli
19	30	85.7	491	1	US-08-283-911-2	Sequence 2, Appli
20	30	85.7	491	1	US-08-245-500A-3	Sequence 3, Appli
21	30	85.7	491	1	US-08-390-546-3	Sequence 3, Appli
22	30	85.7	491	1	US-08-390-479A-3	Sequence 3, Appli
23	30	85.7	491	1	US-08-557-393-3	Sequence 3, Appli
24	30	85.7	491	1	US-08-390-516C-3	Sequence 3, Appli
25	30	85.7	491	1	US-08-390-517A-3	Sequence 3, Appli
26	30	85.7	491	1	US-08-390-515A-3	Sequence 3, Appli
27	30	85.7	491	2	US-08-801-718-3	Sequence 3, Appli
28	29	82.9	9	1	US-08-465-325-138	Sequence 138, App
29	29	82.9	9	4	US-09-115-737-138	Sequence 138, App
30	29	82.9	14	1	US-07-725-331-1	Sequence 1, Appli
31	29	82.9	14	2	US-08-569-188-8	Sequence 8, Appli
32	29	82.9	14	5	PCT-US91-05047-1	Sequence 1, Appli
33	29	82.9	14	5	PCT-US94-07019-8	Sequence 8, Appli
34	29	82.9	16	1	US-07-725-331-4	Sequence 4, Appli
35	29	82.9	16	2	US-08-569-188-1	Sequence 1, Appli
36	29	82.9	16	2	US-08-569-188-2	Sequence 2, Appli
37	29	82.9	16	2	US-08-569-188-10	Sequence 10, Appl
38	29	82.9	16	2	US-08-569-188-11	Sequence 11, Appl
39	29	82.9	16	2	US-08-569-188-12	Sequence 12, Appl
40	29	82.9	16	2	US-08-569-188-13	Sequence 13, Appl
41	29	82.9	16	5	PCT-US91-05047-4	Sequence 4, Appli
42	29	82.9	16	5	PCT-US94-07019-1	Sequence 1, Appli
43	29	82.9	16	5	PCT-US94-07019-2	Sequence 2, Appli
44	29	82.9	16	5	PCT-US94-07019-10	Sequence 10, Appl
45	29	82.9	16	5	PCT-US94-07019-11	Sequence 11, Appl

ALIGNMENTS

```
RESULT   1
US-09-334-601-4
; Sequence 4, Application US/09334601
; Patent No. 6280989
; GENERAL INFORMATION:
;  APPLICANT: Kapitonov, Dmitri
;  APPLICANT: Yu, Robert
;  TITLE OF INVENTION: NOVEL SIALYLTRANSPERASES
;  FILE REFERENCE: VCUIP-6
;  CURRENT APPLICATION NUMBER: US/09/334,601
;  CURRENT FILING DATE: 1999-06-17
;  NUMBER OF SEQ ID NOS: 94
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 414
;   TYPE: PRT
;   ORGANISM: Murinae gen. sp.
US-09-334-601-4


  Query Match             88.6%;  Score 31;  DB 4;  Length 414;
  Best Local Similarity   87.5%;  Pred. No. 2.1e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            | ||||||
Db     395 KFLLKLLK 402


RESULT   2
US-07-725-331-9
; Sequence 9, Application US/07725331
; Patent No. 5294605
;  GENERAL INFORMATION:
;    APPLICANT:  Houghten, Richard
;    APPLICANT:  Blondelle, Sylvie
;    TITLE OF INVENTION:  Amphiphilic Peptide Compositions and
;    TITLE OF INVENTION:  Analogues Thereof
;    NUMBER OF SEQUENCES:  68
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dressler, Goldsmith, Sutker, Shore,
;      ADDRESSEE:  & Milnamow
;      STREET:  180 No. 5294605th Stetson
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.24
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/725,331
;       FILING DATE:
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/554,422
;       FILING DATE:  19-JUL-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gamson, Edward P.
;       REGISTRATION NUMBER:  29,381
;       REFERENCE/DOCKET NUMBER:  421250-80
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  3126165418
;       TELEFAX:  3126165460
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  18 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       OTHER INFORMATION:  May be a C-terminal amide, and/or may
;       OTHER INFORMATION:  be acetylated at N-terminus, Xaa is
;       OTHER INFORMATION:  Met or methionine sulfoxide.
US-07-725-331-9


  Query Match           85.7%;  Score 30;  DB 1;  Length 18;
  Best Local Similarity   87.5%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            ||| |||| 
Db       2 KLLXKLLK 9


RESULT   3
PCT-US91-05047-9
; Sequence 9, Application PC/TUS9105047
;  GENERAL INFORMATION:
;    APPLICANT:  Houghten, Richard
;    APPLICANT:  Blondelle, Sylvie
;    TITLE OF INVENTION:  Amphiphilic Peptide Compositions and
;    TITLE OF INVENTION:  Analogues Thereof
;    NUMBER OF SEQUENCES:  68
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dressler, Goldsmith, Sutker, Shore,
;      ADDRESSEE:  & Milnamow
;      STREET:  180 North Stetson
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/05047
;      FILING DATE:  19910717
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/554,422
;      FILING DATE:  19-JUL-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gamson, Edward P.
;      REGISTRATION NUMBER:  29,381
;      REFERENCE/DOCKET NUMBER:  421250-80
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  3126165418
;      TELEFAX:  3126165460
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  May be a C-terminal amide, and/or may
;      OTHER INFORMATION:  be acetylated at N-terminus, Xaa is
;      OTHER INFORMATION:  Met or methionine sulfoxide.
PCT-US91-05047-9


  Query Match           85.7%;  Score 30;  DB 5;  Length 18;
  Best Local Similarity   87.5%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            ||| |||| 
Db       2 KLLXKLLK 9


RESULT   4
US-08-944-133-13
; Sequence 13, Application US/08944133
; Patent No. 5789542
;  GENERAL INFORMATION:
;    APPLICANT:  McLaughlin, Mark L
;    APPLICANT:  Becker, Calvin L
;    TITLE OF INVENTION:  Amphipathic Peptides
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John H. Runnels
;      STREET:  P. O. Box 2471
;      CITY:  Baton Rouge
;      STATE:  LA
;      COUNTRY:  USA
;      ZIP:  70821-2471
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/944,133
;      FILING DATE:  06-OCT-1997
;      CLASSIFICATION:  5530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/789,077
;      FILING DATE:  03-FEB-1997
;      APPLICATION NUMBER:  US/08/681,075
;      FILING DATE:
;      APPLICATION NUMBER:  US/08/232,525
;      FILING DATE:  22-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Runnels, John H
;      REGISTRATION NUMBER:  33451
;      REFERENCE/DOCKET NUMBER:  Atty File No. 5789542 9301
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  504 387-3221
;      TELEFAX:  504 346-8049
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-944-133-13
```

```
  Query Match             85.7%;  Score 30;  DB 1;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 22;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            || |||||
Db       2 KLALKLLK 9


RESULT   5
US-08-248-839C-47
; Sequence 47, Application US/08248839C
; Patent No. 5843702
;  GENERAL INFORMATION:
;    APPLICANT:  McConnell, David
;    APPLICANT:  Devine, Kevin
;    APPLICANT:  O'Kane, Charles
;    TITLE OF INVENTION:  A Gene Expression System
;    NUMBER OF SEQUENCES:  185
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5843702o No. 5843702disk of No. 5843702th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/248,839C
;      FILING DATE:  25-MAY-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gregg, Valeta A.
;      REGISTRATION NUMBER:  35,127
;      REFERENCE/DOCKET NUMBER:  3614.214-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  86 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Protein
US-08-248-839C-47


  Query Match             85.7%;  Score 30;  DB 2;  Length 86;
  Best Local Similarity   75.0%;  Pred. No. 76;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KLLLKLLK 8
            |::|||||
Db      49 KIVLKLLK 56


RESULT   6
US-07-903-103-4
; Sequence 4, Application US/07903103
; Patent No. 5411860
;  GENERAL INFORMATION:
;    APPLICANT:  VOGELSTEIN, BERT
;    APPLICANT:  KINZLER, KENNETH
;    TITLE OF INVENTION:  AMPLIFICATION OF HUMAN MDM2 GENE IN
;    TITLE OF INVENTION:  HUMAN TUMORS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER, BIRCH, MCKIE AND BECKETT
;      STREET:  1001 G ST., N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001-4597
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/903,103
;      FILING DATE:  19920623
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/867.840
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KAGAN, SARAH A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.40148
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  489 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-903-103-4


  Query Match             85.7%;  Score 30;  DB 1;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
            |||||||
Db      33 LLLKLLK 39


RESULT   7
US-08-044-619A-4
; Sequence 4, Application US/08044619A
; Patent No. 5420263
;  GENERAL INFORMATION:
;    APPLICANT:  THE JOHNS HOPKINS UNIVERSITY
;    APPLICANT:  720 RUTLAND AVENUE, BALTIMORE, MARYLAND 21205 USA
;    TITLE OF INVENTION:  AMPLIFICATION OF HUMAN MDM2 GENE IN
;    TITLE OF INVENTION:  HUMAN TUMORS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER, BIRCH, MCKIE AND BECKETT
;      STREET:  1001 G ST., N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001-4597
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/044,619A
;      FILING DATE:  07-APR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/903,103
;      FILING DATE:  23-JUN-1992
```

```
;       APPLICATION NUMBER:  US 07/867.840
;       FILING DATE:  07-APR-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  KAGAN, SARAH A.
;       REGISTRATION NUMBER:  32,141
;       REFERENCE/DOCKET NUMBER:  01107.40148
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-508-9100
;       TELEFAX:  202-508-9299
;       TELEX:  197430 BBMB UT
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  489 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-044-619A-4


  Query Match             85.7%;  Score 30;  DB 1;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT   8
US-08-283-911-4
; Sequence 4, Application US/08283911
; Patent No. 5519118
;  GENERAL INFORMATION:
;    APPLICANT:  VOGELSTEIN, BERT
;    APPLICANT:  KINZLER, KENNETH
;    TITLE OF INVENTION:  AMPLIFICATION OF HUMAN MDM2 GENE IN
;    TITLE OF INVENTION:  HUMAN TUMORS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER, BIRCH, MCKIE AND BECKETT
;      STREET:  1001 G ST., N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001-4597
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/283,911
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/903,103
;      FILING DATE:  23-JUN-1992
;      APPLICATION NUMBER:  US 07/867.840
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KAGAN, SARAH A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.40148
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  489 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-283-911-4


  Query Match             85.7%;  Score 30;  DB 1;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT   9
US-08-245-500A-5
; Sequence 5, Application US/08245500A
; Patent No. 5550023
;  GENERAL INFORMATION:
;    APPLICANT:  BURRELL, MARILEE
;    APPLICANT:  HILL, DAVID E.
;    APPLICANT:  KINZLER, KENNETH W.
;    APPLICANT:  VOGELSTEIN, BERT
;    TITLE OF INVENTION:  AMPLIFICATION OF HUMAN MDM2 GENE IN
;    TITLE OF INVENTION:  HUMAN TUMORS
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER, BIRCH, MCKIE AND BECKETT
;      STREET:  1001 G STREET, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/245,500A
;      FILING DATE:  07-APR-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KAGAN, SARAH A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.42798
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  489 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-245-500A-5


  Query Match             85.7%;  Score 30;  DB 1;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT  10
US-08-390-546-5
; Sequence 5, Application US/08390546
; Patent No. 5606044
;  GENERAL INFORMATION:
;    APPLICANT:  BURRELL, MARILEE
;    APPLICANT:  HILL, DAVID E.
```

```
;    APPLICANT:  KINZLER, KENNETH W.
;    APPLICANT:  VOGELSTEIN, BERT
;    TITLE OF INVENTION:  AMPLIFICATION OF HUMAN MDM2 GENE IN
;    TITLE OF INVENTION:  HUMAN TUMORS
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER, BIRCH, MCKIE AND BECKETT
;      STREET:  1001 G STREET, N°W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/390,546
;      FILING DATE:  07-APR-1993
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KAGAN, SARAH A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.42798
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  489 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-390-546-5


  Query Match             85.7%;  Score 30;  DB 1;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT  11
US-08-390-479A-5
; Sequence 5, Application US/08390479A
; Patent No. 5618921
;  GENERAL INFORMATION:
;    APPLICANT:  BURRELL, MARILEE
;    APPLICANT:  HILL, DAVID E.
;    APPLICANT:  KINZLER, KENNETH W.
;    APPLICANT:  VOGELSTEIN, BERT
;    TITLE OF INVENTION:  AMPLIFICATION OF HUMAN MDM2 GENE IN
;    TITLE OF INVENTION:  HUMAN TUMORS
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER & WITCOFF, LTD.
;      STREET:  1001 G STREET, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/390,479A
;      FILING DATE:  02-FEB-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KAGAN, SARAH A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.48992
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  489 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-390-479A-5


  Query Match             85.7%;  Score 30;  DB 1;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLLKLLK 8
           |||||||
Db      33 LLLKLLK 39


RESULT  12
US-08-557-393-5
; Sequence 5, Application US/08557393
; Patent No. 5702903
;  GENERAL INFORMATION:
;    APPLICANT:  BURRELL, MARILEE
;    APPLICANT:  HILL, DAVID E.
;    APPLICANT:  KINZLER, KENNETH W.
;    APPLICANT:  VOGELSTEIN, BERT
;    TITLE OF INVENTION:  AMPLIFICATION OF HUMAN MDM2 GENE IN
;    TITLE OF INVENTION:  HUMAN TUMORS
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER, BIRCH, MCKIE AND BECKETT
;      STREET:  1001 G STREET, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/557,393
;      FILING DATE:  13-NOV-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/245,500
;      FILING DATE:  18-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KAGAN, SARAH A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.42798
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  489 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
```

US-08-557-393-5

Query Match 85.7%; Score 30; DB 1; Length 489;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 LLLKLLK 8
          |||||||
Db     33 LLLKLLK 39
```

RESULT 13
US-08-390-516C-5
; Sequence 5, Application US/08390516C
; Patent No. 5708136
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,516C
; FILING DATE: 07-APR-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 489 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-390-516C-5

Query Match 85.7%; Score 30; DB 1; Length 489;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 LLLKLLK 8
          |||||||
Db     33 LLLKLLK 39
```

RESULT 14
US-08-390-517A-5
; Sequence 5, Application US/08390517A
; Patent No. 5736338
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,517A
; FILING DATE: 07-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 489 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-390-517A-5

Query Match 85.7%; Score 30; DB 1; Length 489;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 LLLKLLK 8
          |||||||
Db     33 LLLKLLK 39
```

RESULT 15
US-08-390-515A-5
; Sequence 5, Application US/08390515A
; Patent No. 5756455
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,515A

; FILING DATE: 07-APR-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 489 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-390-515A-5

Query Match 85.7%; Score 30; DB 1; Length 489;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 LLLKLLK 8
            |||||||
Db      33 LLLKLLK 39
```

Search completed: June 17, 2002, 12:42:05
Job time: 225 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:41:23 ; Search time 94.14 Seconds
(without alignments)
16.518 Million cell updates/sec

Title: US-09-367-714A-92
Perfect score: 70
Sequence: 1 CKLLLKLLLKLLKC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	70	100.0	14	19	AAW82854	Antipathogenic pep
2	70	100.0	77	19	AAW82858	Antipathogenic pep
3	64	91.4	14	19	AAW82855	Antipathogenic pep
4	64	91.4	77	19	AAW82859	Antipathogenic pep
5	61	87.1	13	18	AAW35231	Diastereomer pepti
6	61	87.1	13	21	AAB17482	Antipathogenic pep
7	55	78.6	13	18	AAW35232	Diastereomer pepti
8	55	78.6	13	21	AAB17484	Antipathogenic pep
9	52	74.3	12	18	AAW35149	Leu/Lys diastereom
10	52	74.3	12	18	AAW35152	Leu/Lys diastereom
11	52	74.3	12	19	AAW82847	Antipathogenic pep
12	52	74.3	12	19	AAW82850	Antipathogenic pep
13	52	74.3	12	19	AAW82856	Antipathogenic pep
14	52	74.3	12	21	AAB17413	Antipathogenic pep
15	52	74.3	12	21	AAB17416	Antipathogenic pep
16	52	74.3	12	21	AAB17483	Antipathogenic pep
17	52	74.3	12	21	AAB17485	Antipathogenic pep
18	52	74.3	37	19	AAW77378	Lytic peptide with
19	46	65.7	12	18	AAW35150	Leu/Lys diastereom
20	46	65.7	12	18	AAW35153	Leu/Lys diastereom
21	46	65.7	12	18	AAW35169	Leu/Lys diastereom
22	46	65.7	12	18	AAW35170	Leu/Lys diastereom
23	46	65.7	12	18	AAW35171	Leu/Lys diastereom
24	46	65.7	12	19	AAW82848	Antipathogenic pep
25	46	65.7	12	19	AAW82851	Antipathogenic pep
26	46	65.7	12	19	AAW82857	Antipathogenic pep
27	46	65.7	12	19	AAW82885	Antipathogenic pep
28	46	65.7	12	19	AAW82886	Antipathogenic pep
29	46	65.7	12	19	AAW82887	Antipathogenic pep
30	46	65.7	12	21	AAB17414	Antipathogenic pep
31	46	65.7	12	21	AAB17417	Antipathogenic pep
32	46	65.7	12	21	AAB17420	Antipathogenic pep
33	46	65.7	12	21	AAB17421	Antipathogenic pep
34	46	65.7	12	21	AAB17422	Antipathogenic pep
35	46	65.7	21	22	AAU03187	Membrane active sy
36	46	65.7	21	22	AAB60066	KL3 membrane activ
37	46	65.7	153	20	AAY29393	Sperm whale myoglo
38	45	64.3	20	19	AAW65895	Amphipathic peptid
39	45	64.3	20	21	AAY49957	Amphipathic peptid
40	43	61.4	73	21	AAB52057	Human secreted pro
41	42	60.0	1300	22	ABB64083	Drosophila melanog
42	41.5	59.3	15	19	AAW77384	Lytic peptide with
43	41	58.6	28	10	AAP91335	Amino acid sequenc
44	40	57.1	14	16	AAR67795	Antimicrobial olig
45	40	57.1	14	21	AAB17122	Calmodulin antagon

ALIGNMENTS

RESULT 1
AAW82854
ID AAW82854 standard; peptide; 14 AA.
XX
AC AAW82854;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX

PS Claim 14; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 70; DB 19; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00074;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKLLLKLLLKLLKC 14
||||||||||||||
Db 1 cklllklllkllkc 14

RESULT 2
AAW82858
ID AAW82858 standard; peptide; 77 AA.
XX
AC AAW82858;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 17; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 77 AA;

Query Match 100.0%; Score 70; DB 19; Length 77;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKLLLKLLLKLLKC 14
||||||||||||||
Db 13 cklllklllkllkc 26

RESULT 3
AAW82855
ID AAW82855 standard; peptide; 14 AA.
XX
AC AAW82855;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 14; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which

CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 14 AA;

Query Match 91.4%; Score 64; DB 19; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.0053;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKLLLKLLLKLLKC 14
|||||||| ||||||
Db 1 cklllklklkllkc 14

RESULT 4
AAW82859
ID AAW82859 standard; peptide; 77 AA.
XX
AC AAW82859;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 17; Page 107; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 77 AA;

Query Match 91.4%; Score 64; DB 19; Length 77;
Best Local Similarity 92.9%; Pred. No. 0.025;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKLLLKLLLKLLKC 14
|||||||| ||||||
Db 13 cklllklklkllkc 26

RESULT 5
AAW35231
ID AAW35231 standard; peptide; 13 AA.
XX
AC AAW35231;
XX
DT 14-APR-1998 (first entry)
XX
DE Diastereomer peptide [D]-L3,4,8,10-K4L8C.
XX
KW Diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW Honey bee venom; pardaxin; cytolytic activity; cancer;
KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW agricultural pesticide; cell wall lysis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 3
FT /note= "D-form residue"
FT Misc-difference 4
FT /note= "D-form residue"
FT Misc-difference 8
FT /note= "D-form residue"
FT Misc-difference 10
FT /note= "D-form residue"
XX
PN WO9731019-A2.
XX
PD 28-AUG-1997.
XX
PF 20-FEB-1997; 97WO-IL00066.
XX
PR 22-FEB-1996; 96IL-0117223.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer
XX
PS Example 7; Page 49; 80pp; English.
XX
CC This sequence represents a diastereomer peptide of the invention. This
CC sequence is used in a "bundle sequence", which is created by binding 5
CC copies of this sequence to peptide 23 (see AAW35149). The peptides of
CC the invention have: (a) cytolytic activity on pathogenic cells (pathogens
CC and malignant cells not naturally present in the body); but (b) no
CC haemolytic activity, or such activity only at a concentration
CC significantly higher than that at which they lyse pathogens. The

CC peptides, their complexes and mixtures are used to treat infections
CC (caused by bacteria, fungi, protozoa, mycoplasma or viruses) or cancer,
CC in human and veterinary medicine. Also, they can be used as preservatives
CC for food, cosmetics and agricultural produce, or as agricultural
CC pesticides. The absence of haemolytic activity (associated with
CC disturbance of alpha-helical structures) means that the peptides have few
CC if any toxic effects, and those that include D-aa will have increased
CC resistance to proteolytic degradation. Non-haemolytic, cytotoxic random
CC copolymers of pardaxin, each has a specific spectrum of activity,
CC allowing selection of agents for particular applications. Since these
CC random copolymers induce total lysis of bacterial cell walls, resistance
CC to them is unlikely to develop.
XX
SQ Sequence 13 AA;

Query Match 87.1%; Score 61; DB 18; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLKC 14
||||||||||||||
Db 1 klllklllkllkc 13

RESULT 6
AAB17482
ID AAB17482 standard; Peptide; 13 AA.
XX
AC AAB17482;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:586.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 401; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 13 AA;

Query Match 87.1%; Score 61; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLKC 14
||||||||||||||
Db 1 klllklllkllkc 13

RESULT 7
AAW35232
ID AAW35232 standard; peptide; 13 AA.
XX
AC AAW35232;
XX
DT 14-APR-1998 (first entry)
XX
DE Diastereomer peptide [D]-L3,4,8,10-K5L7C.
XX
KW Diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW Honey bee venom; pardaxin; cytolytic activity; cancer;
KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW agricultural pesticide; cell wall lysis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 3
FT /note= "D-form residue"
FT Misc-difference 4
FT /note= "D-form residue"
FT Misc-difference 8
FT /note= "D-form residue"
FT Misc-difference 10
FT /note= "D-form residue"
XX
PN WO9731019-A2.
XX
PD 28-AUG-1997.
XX
PF 20-FEB-1997; 97WO-IL00066.
XX
PR 22-FEB-1996; 96IL-0117223.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer
XX
PS Example 7; Page 50; 80pp; English.
XX
CC This sequence represents a diastereomer peptide of the invention. This
CC sequence is used in a "bundle sequence", which is created by binding 5

CC copies of this sequence to peptide 26 (see AAW35152). The peptides of
CC the invention have: (a) cytolytic activity on pathogenic cells (pathogens
CC and malignant cells not naturally present in the body); but (b) no
CC haemolytic activity, or such activity only at a concentration
CC significantly higher than that at which they lyse pathogens. The
CC peptides, their complexes and mixtures are used to treat infections
CC (caused by bacteria, fungi, protozoa, mycoplasma or viruses) or cancer,
CC in human and veterinary medicine. Also, they can be used as preservatives
CC for food, cosmetics and agricultural produce, or as agricultural
CC pesticides. The absence of haemolytic activity (associated with
CC disturbance of alpha-helical structures) means that the peptides have few
CC if any toxic effects, and those that include D-aa will have increased
CC resistance to proteolytic degradation. Non-haemolytic, cytotoxic random
CC copolymers of pardaxin, each has a specific spectrum of activity,
CC allowing selection of agents for particular applications. Since these
CC random copolymers induce total lysis of bacterial cell walls, resistance
CC to them is unlikely to develop.
XX
SQ Sequence 13 AA;

Query Match 78.6%; Score 55; DB 18; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.094;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLKC 14
|||||| ||||||
Db 1 klllklklkllkc 13

RESULT 8
AAB17484
ID AAB17484 standard; Peptide; 13 AA.
XX
AC AAB17484;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:588.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 401; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 13 AA;

Query Match 78.6%; Score 55; DB 21; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.094;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLKC 14
|||||| ||||||
Db 1 klllklklkllkc 13

RESULT 9
AAW35149
ID AAW35149 standard; peptide; 12 AA.
XX
AC AAW35149;
XX
DT 14-APR-1998 (first entry)
XX
DE Leu/Lys diastereomer peptide [D]-L3,4,8,10-K4L8.
XX
KW Leu/Lys diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW Honey bee venom; pardaxin; cytolytic activity; cancer;
KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW agricultural pesticide; cell wall lysis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 3
FT /note= "D-form residue"
FT Misc-difference 4
FT /note= "D-form residue"
FT Misc-difference 8
FT /note= "D-form residue"
FT Misc-difference 10
FT /note= "D-form residue"
FT Modified-site 12
FT /note= "C-terminal amide"
XX
PN WO9731019-A2.
XX
PD 28-AUG-1997.
XX
PF 20-FEB-1997; 97WO-IL00066.
XX
PR 22-FEB-1996; 96IL-0117223.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1997-435088/40.
XX
PT Peptide(s) having selective cytolytic activity - against pathogens

```
PT   and malignant cells, but no haemolytic activity, used for treating
PT   infections and cancer
XX
PS   Claim 21; Page 39; 80pp; English.
XX
CC   This sequence represents a Leu/Lys diastereomer peptide of the
CC   invention. The peptides of the invention have: (a) cytolytic activity on
CC   pathogenic cells (pathogens and malignant cells not naturally present in
CC   the body); but (b) no haemolytic activity, or such activity only at a
CC   concentration significantly higher than that at which they lyse
CC   pathogens. The peptides, their complexes and mixtures are used to treat
CC   infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC   or cancer, in human and veterinary medicine. Also, they can be used as
CC   preservatives for food, cosmetics and agricultural produce, or as
CC   agricultural pesticides. The absence of haemolytic activity (associated
CC   with disturbance of alpha-helical structures) means that the peptides
CC   have few if any toxic effects, and those that include D-aa will have
CC   increased resistance to proteolytic degradation. Non-haemolytic,
CC   cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC   activity, allowing selection of agents for particular applications. Since
CC   these random copolymers induce total lysis of bacterial cell walls,
CC   resistance to them is unlikely to develop.
XX
SQ   Sequence   12 AA;


  Query Match             74.3%;  Score 52;  DB 18;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 KLLLKLLLKLLK 13
          ||||||||||||
Db      1 klllklllkllk 12


RESULT  10
AAW35152
ID   AAW35152 standard; peptide; 12 AA.
XX
AC   AAW35152;
XX
DT   14-APR-1998  (first entry)
XX
DE   Leu/Lys diastereomer peptide [D]-K1,5,9,12L2,6,7,11-K4L8.
XX
KW   Leu/Lys diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW   Honey bee venom; pardaxin; cytolytic activity; cancer;
KW   non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW   agricultural pesticide; cell wall lysis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 2
FT                   /note= "D-form residue"
FT   Misc-difference 5
FT                   /note= "D-form residue"
FT   Misc-difference 6
FT                   /note= "D-form residue"
FT   Misc-difference 7
FT                   /note= "D-form residue"
FT   Misc-difference 9
FT                   /note= "D-form residue"
FT   Misc-difference 11
FT                   /note= "D-form residue"
FT   Misc-difference 12
FT                   /note= "D-form residue"
FT   Modified-site   12
FT                   /note= "C-terminal amide"
XX
PN   WO9731019-A2.
XX
PD   28-AUG-1997.
XX
PF   20-FEB-1997;   97WO-IL00066.
XX
PR   22-FEB-1996;   96IL-0117223.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Oren Z,  Shai Y;
XX
DR   WPI; 1997-435088/40.
XX
PT   Peptide(s) having selective cytolytic activity - against pathogens
PT   and malignant cells, but no haemolytic activity, used for treating
PT   infections and cancer
XX
PS   Claim 21; Page 40; 80pp; English.
XX
CC   This sequence represents a Leu/Lys diastereomer peptide of the
CC   invention. The peptides of the invention have: (a) cytolytic activity on
CC   pathogenic cells (pathogens and malignant cells not naturally present in
CC   the body); but (b) no haemolytic activity, or such activity only at a
CC   concentration significantly higher than that at which they lyse
CC   pathogens. The peptides, their complexes and mixtures are used to treat
CC   infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC   or cancer, in human and veterinary medicine. Also, they can be used as
CC   preservatives for food, cosmetics and agricultural produce, or as
CC   agricultural pesticides. The absence of haemolytic activity (associated
CC   with disturbance of alpha-helical structures) means that the peptides
CC   have few if any toxic effects, and those that include D-aa will have
CC   increased resistance to proteolytic degradation. Non-haemolytic,
CC   cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC   activity, allowing selection of agents for particular applications. Since
CC   these random copolymers induce total lysis of bacterial cell walls,
CC   resistance to them is unlikely to develop.
XX
SQ   Sequence   12 AA;


  Query Match             74.3%;  Score 52;  DB 18;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 KLLLKLLLKLLK 13
          ||||||||||||
Db      1 klllklllkllk 12


RESULT  11
AAW82847
ID   AAW82847 standard; peptide; 12 AA.
XX
AC   AAW82847;
XX
DT   19-MAY-1999  (first entry)
XX
DE   Antipathogenic peptide.
XX
KW   Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW   cancer; infection; disinfectant; contact lens wetting solution;
KW   preservative; pesticide; fungicide; bactericide.
XX
OS   Synthetic.
XX
PN   WO9837090-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;   98WO-IL00081.
XX
PR   20-FEB-1997;   97WO-IL00066.
XX
```

PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 12; Page 105; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 74.3%; Score 52; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLK 13
||||||||||||
Db 1 klllklllkllk 12

RESULT 12
AAW82850
ID AAW82850 standard; peptide; 12 AA.
XX
AC AAW82850;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX
PS Claim 13; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 74.3%; Score 52; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLK 13
||||||||||||
Db 1 klllklllkllk 12

RESULT 13
AAW82856
ID AAW82856 standard; peptide; 12 AA.
XX
AC AAW82856;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antipathogenic peptide.
XX
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
XX
OS Synthetic.
XX
PN WO9837090-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-IL00081.
XX
PR 20-FEB-1997; 97WO-IL00066.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Oren Z, Shai Y;
XX
DR WPI; 1998-594464/50.
XX
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
XX

PS Claim 14; Page 106; 126pp; English.
XX
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
XX
SQ Sequence 12 AA;

Query Match 74.3%; Score 52; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLK 13
||||||||||||
Db 1 klllklllkllk 12

RESULT 14
AAB17413
ID AAB17413 standard; Peptide; 12 AA.
XX
AC AAB17413;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:517.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX

PS Claim 39; Page 378; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 74.3%; Score 52; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLK 13
||||||||||||
Db 1 klllklllkllk 12

RESULT 15
AAB17416
ID AAB17416 standard; Peptide; 12 AA.
XX
AC AAB17416;
XX
DT 31-OCT-2000 (first entry)
XX
DE Antipathogenic peptide sequence SEQ ID NO:520.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 379; 608pp; English.

XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

Query Match 74.3%; Score 52; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 KLLLKLLLKLLK 13
          ||||||||||||
Db      1 klllklllkllk 12
```

Search completed: June 17, 2002, 12:41:23
Job time: 298 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:43:01 ; Search time 46.42 Seconds
(without alignments)
28.980 Million cell updates/sec

Title: US-09-367-714A-92
Perfect score: 70
Sequence: 1 CKLLLKLLLKLLKC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	61.4	38	2	B85632	hypothetical prote
2	40	57.1	169	2	G72096	hypothetical prote
3	40	57.1	169	2	E86525	hypothetical prote
4	40	57.1	314	2	S50368	probable membrane
5	40	57.1	1880	2	T18531	tractin - medicina
6	40	57.1	2513	2	G96536	hypothetical prote
7	39	55.7	641	2	G85043	hypothetical prote
8	38	54.3	30	2	B44314	intracisternal A p
9	38	54.3	219	2	A35650	Sur protein - chic
10	38	54.3	754	2	A85043	probable LRR recep
11	38	54.3	1107	2	T20578	hypothetical prote
12	38	54.3	1119	2	T20577	hypothetical prote
13	37	52.9	83	2	AH1734	hypothetical prote
14	37	52.9	116	2	T41597	hypothetical prote
15	37	52.9	137	2	A96914	uncharacterized pr
16	37	52.9	161	2	T48285	hypothetical prote
17	37	52.9	230	2	JC5418	somatolactin precu
18	37	52.9	238	2	E71375	probable ABC trans
19	37	52.9	255	2	A60637	merozoite antigen
20	37	52.9	275	2	H71690	hypothetical prote
21	37	52.9	465	2	T30155	hypothetical prote
22	37	52.9	481	2	T22406	hypothetical prote
23	37	52.9	935	2	S63261	SEC21 protein - ye
24	37	52.9	1030	2	T16114	hypothetical prote
25	36	51.4	143	2	S03747	small membrane pro
26	36	51.4	230	2	JC2582	somatolactin - Sen
27	36	51.4	231	2	A35793	somatolactin - Jap
28	36	51.4	238	2	A31417	prolactin-related
29	36	51.4	375	2	C64216	hypothetical prote
30	36	51.4	399	2	T20204	hypothetical prote
31	36	51.4	438	2	AH0031	proton glutamate s
32	36	51.4	568	2	T17308	hypothetical prote
33	36	51.4	1661	2	T31330	head-activator bin
34	36	51.4	1896	2	T01490	hypothetical prote
35	35	50.0	39	2	G85837	hypothetical prote
36	35	50.0	125	2	I50498	urotensin II-alpha
37	35	50.0	164	2	B72717	hypothetical prote
38	35	50.0	167	2	H84493	probable replicati
39	35	50.0	179	2	E75204	hypothetical prote
40	35	50.0	219	2	C84647	probable synaptobr
41	35	50.0	235	2	A49762	somatolactin precu
42	35	50.0	318	2	C81386	probable integral
43	35	50.0	319	2	A70102	conserved hypothet
44	35	50.0	373	2	T18924	hypothetical prote
45	35	50.0	394	2	T19116	hypothetical prote

ALIGNMENTS

RESULT 1
B85632
hypothetical protein Z1386 [imported] - Escherichia coli (strain O157:H7, substrain E
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: B85632
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85632
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-38 <STO>
A;Cross-references: GB:AE005174; NID:g12514232; PIDN:AAG55518.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z1386

Query Match 61.4%; Score 43; DB 2; Length 38;
Best Local Similarity 42.9%; Pred. No. 2.1;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 CKLLLKLLLKLLKC 14
           |::|: :|||:: |
Db       6 CEILVNILLKMIGC 19
```

RESULT 2
G72096
hypothetical protein CP0481 [imported] - Chlamydophila pneumoniae (strains CWL029 and
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: G72096; C81573
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: G72096
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <ARN>
A;Cross-references: GB:AE001613; GB:AE001363; NID:g4376550; PIDN:AAD18426.1; PID:g437
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255
A;Accession: C81573

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <REA>
A;Cross-references: GB:AE002209; GB:AE002161; NID:g7189393; PIDN:AAF38311.1; PID:g718939
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CPn0277; CP0481
C;Superfamily: Chlamydia pneumoniae hypothetical protein CPn0277


  Query Match             57.1%;  Score 40;  DB 2;  Length 169;
  Best Local Similarity   50.0%;  Pred. No. 23;
  Matches    6;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLLKLLLKLLKC 14
           |:|::::||:|| 
Db      94 LVLQVIIKLIKC 105


RESULT   3
E86525
hypothetical protein CPj0277 [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: E86525
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: E86525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <STO>
A;Cross-references: GB:BA000008; NID:g8978651; PIDN:BAA98487.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: CPj0277
C;Superfamily: Chlamydia pneumoniae hypothetical protein CPn0277


  Query Match             57.1%;  Score 40;  DB 2;  Length 169;
  Best Local Similarity   50.0%;  Pred. No. 23;
  Matches    6;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLLKLLLKLLKC 14
           |:|::::||:|| 
Db      94 LVLQVIIKLIKC 105


RESULT   4
S50368
probable membrane protein YLR283w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L8003.9
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C;Accession: S50368
R;Pauley, A.
submitted to the EMBL Data Library, November 1994
A;Description: The sequence of S. cerevisiae cosmid 8003.
A;Reference number: S50366
A;Accession: S50368
A;Molecule type: DNA
A;Residues: 1-314 <PAU>
A;Cross-references: EMBL:U17243; NID:g596030; PID:g596039; GSPDB:GN00012; MIPS:YLR283w
C;Genetics:
A;Gene: MIPS:YLR283w
A;Map position: 12R
C;Keywords: transmembrane protein
F;263-279/Domain: transmembrane #status predicted <TMM>


  Query Match             57.1%;  Score 40;  DB 2;  Length 314;
  Best Local Similarity   58.3%;  Pred. No. 38;
  Matches    7;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLL 12
           ||::: |||:|| 
Db     122 CKIIIALLLQLL 133


RESULT   5
T18531
tractin - medicinal leech
C;Species: Hirudo medicinalis (medicinal leech)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C;Accession: T18531
R;Huang, Y.; Jellies, J.; Johansen, K.M.; Johansen, J.
J. Cell Biol. 138, 143-157, 1997
A;Title: Differential glycosylation of Tractin and LeechCAM, two novel Ig-superfamily
A;Reference number: Z18951; MUID:97362067
A;Accession: T18531
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1880 <HUA>
A;Cross-references: EMBL:U92813; NID:g2275259; PID:g2275260; PIDN:AAC47654.1


  Query Match             57.1%;  Score 40;  DB 2;  Length 1880;
  Best Local Similarity   64.3%;  Pred. No. 1.7e+02;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLLKC 14
           | ||| ||| :| | 
Db    1758 CLLLLLLLLLILLC 1771


RESULT   6
G96536
hypothetical protein F2J10.9 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: G96536
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: G96536
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2513 <STO>
A;Cross-references: GB:AE005173; NID:g8569097; PIDN:AAF76442.1; GSPDB:GN00141
C;Genetics:
A;Gene: F2J10.9
A;Map position: 1


  Query Match             57.1%;  Score 40;  DB 2;  Length 2513;
  Best Local Similarity   90.9%;  Pred. No. 2.1e+02;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 LLLKLLLKLLK 13
           ||| ||||||| 
Db     797 LLLGLLLKLLK 807


RESULT   7
G85043
hypothetical protein AT4g03450 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
```

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: G85043
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488
A;Accession: G85043
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <STO>
A;Cross-references: GB:NC_001268; NID:g7270215; PIDN:CAB77830.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g03450
A;Map position: 4

Query Match 55.7%; Score 39; DB 2; Length 641;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LKLLLKLLKC 14
:|: ||||||
Db 304 VKIFLKLLKC 313

RESULT 8
B44314
intracisternal A particle GAG protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
C;Accession: B44314
R;Brigle, K.E.; Westin, E.H.; Houghton, M.T.; Goldman, I.D.
J. Biol. Chem. 267, 22351-22355, 1992
A;Title: Insertion of an intracisternal A particle within the 5'-regulatory region of a
ith increased protein expression.
A;Reference number: A44314; MUID:93054523
A;Accession: B44314
A;Status: preliminary
A;Molecule type: DNA; protein
A;Residues: 1-30 <BRI>
A;Experimental source: L1210 leukemia cells LL1 subline
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:117028, NCBIP:117029)
C;Superfamily: AIDS-related virus gag polyprotein

Query Match 54.3%; Score 38; DB 2; Length 30;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKLLLKLLLKLL 12
| |:||| ||||
Db 19 CYLVLKLGLKLL 30

RESULT 9
A35650
Sur protein - chicken
C;Species: Gallus gallus (chicken)
C;Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 21-Jul-2000
C;Accession: A35650
R;Dorai, T.; Wang, L.H.
Mol. Cell. Biol. 10, 4068-4079, 1990
A;Title: An alternative non-tyrosine protein kinase product of the c-src gene in chicken
A;Reference number: A35650; MUID:90318371
A;Accession: A35650
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-219 <DOR>
A;Cross-references: GB:M57290; NID:g212703; PIDN:AAA49076.1; PID:g212704

Query Match 54.3%; Score 38; DB 2; Length 219;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CKLLLKLLLKLLKC 14
|:|::| | ||:|
Db 73 CELMIKCGLVLLQC 86

RESULT 10
A85043
probable LRR receptor-like protein kinase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: A85043
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488
A;Accession: A85043
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-754 <STO>
A;Cross-references: GB:NC_001268; NID:g7270209; PIDN:CAB77824.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g03390
A;Map position: 4

Query Match 54.3%; Score 38; DB 2; Length 754;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CKLLLKLLLKLL 12
| ||| ||| ||
Db 9 CLLLLPLLLSLL 20

RESULT 11
T20578
hypothetical protein F08B12.3b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T20578
R;Dobson, R.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19295
A;Accession: T20578
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1107 <WIL>
A;Cross-references: EMBL:Z68104; PIDN:CAA92116.1; GSPDB:GN00028; CESP:F08B12.3b
A;Experimental source: clone F08B12
C;Genetics:
A;Gene: CESP:F08B12.3b
A;Map position: X
A;Introns: 38/2; 64/1; 112/2; 148/3; 173/3; 201/3; 348/3; 392/2; 452/1; 488/2; 538/2;

Query Match 54.3%; Score 38; DB 2; Length 1107;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CKLLLKLLLKLLKC 14
| :|:||| :| |
Db 70 CNVLIKLLTCILYC 83

RESULT 12
T20577
hypothetical protein F08B12.3a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T20577

R;Dobson, R.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19295
A;Accession: T20577
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1119 <WIL>
A;Cross-references: EMBL:Z68104; PIDN:CAA92115.1; GSPDB:GN00028; CESP:F08B12.3a
A;Experimental source: clone F08B12
C;Genetics:
A;Gene: CESP:F08B12.3a
A;Map position: X
A;Introns: 38/2; 64/1; 112/2; 148/3; 173/3; 201/3; 348/3; 392/2; 452/1; 488/2; 538/2; 80

Query Match 54.3%; Score 38; DB 2; Length 1119;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CKLLLKLLLKLLKC 14
| :|:||| :| |
Db 70 CNVLIKLLTCILYC 83

RESULT 13
AH1734
hypothetical protein lin2421 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AH1734
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1734
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-83 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97648.1; PID:g16414943; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2421

Query Match 52.9%; Score 37; DB 2; Length 83;
Best Local Similarity 63.6%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LLKLLLKLLKC 14
:| || |:|||
Db 50 MLPLLAKILKC 60

RESULT 14
T41597
hypothetical protein SPCC757.06 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T41597
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z22003
A;Accession: T41597
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-116 <LYN>
A;Cross-references: EMBL:AL031825; PIDN:CAA21231.1; GSPDB:GN00068; SPDB:SPCC757.06
A;Experimental source: strain 972h-; cosmid c757
C;Genetics:
A;Gene: SPDB:SPCC757.06
A;Map position: 3
A;Introns: 10/3; 48/3

Query Match 52.9%; Score 37; DB 2; Length 116;
Best Local Similarity 46.2%; Pred. No. 51;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLKC 14
| |:||::|: |
Db 82 KFLIKLIVKMTDC 94

RESULT 15
A96914
uncharacterized protein, Yje/RRF2 family [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: A96914
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: A96914
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-137 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK78100.1; PID:g15022941; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0115

Query Match 52.9%; Score 37; DB 2; Length 137;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KLLLKLLLKLLK 13
: ||||| ||:|
Db 40 RFLLKLLRKLIK 51

Search completed: June 17, 2002, 12:43:02
Job time: 257 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:44:47 ; Search time 21.35 Seconds
(without alignments)
25.390 Million cell updates/sec

Title: US-09-367-714A-92
Perfect score: 70
Sequence: 1 CKLLLKLLLKLLKC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	57.1	229	1	SOML_TETMU	Q9i9h4 tetraodon m
2	40	57.1	231	1	SOM1_SPAAU	P54863 sparus aura
3	40	57.1	231	1	SOM2_SPAAU	P79894 sparus aura
4	40	57.1	231	1	SOML_SCIOC	Q9ygk7 sciaenops o
5	40	57.1	231	1	SOML_SIGGU	Q9pwg4 siganus gut
6	37	52.9	230	1	SOML_CARAU	P79697 carassius a
7	37	52.9	238	1	Y035_TREPA	O83078 treponema p
8	37	52.9	255	1	LP61_EIMTE	P15714 eimeria ten
9	37	52.9	935	1	COPG_YEAST	P32074 saccharomyc
10	36	51.4	143	1	EAG_BACSU	P06630 bacillus su
11	36	51.4	230	1	SOML_HIPHI	P45641 hippoglossu
12	36	51.4	230	1	SOML_SOLSE	P45642 solea seneg
13	36	51.4	231	1	SOML_PAROL	P20362 paralichthy
14	36	51.4	238	1	PRR1_BOVIN	P05402 bos taurus
15	36	51.4	375	1	Y147_MYCGE	P47393 mycoplasma
16	36	51.4	2353	1	CCAH_HUMAN	O95180 homo sapien
17	35	50.0	125	1	UR2A_CYPCA	P04560 cyprinus ca
18	35	50.0	232	1	SOML_ACITR	O93262 acipenser t
19	35	50.0	235	1	SOML_GADMO	P21919 gadus morhu
20	35	50.0	242	1	MTGA_KLEPN	Q48465 klebsiella
21	35	50.0	492	1	SYYM_YEAST	P48527 saccharomyc
22	35	50.0	1033	1	YDK9_SCHPO	P87115 schizosacch
23	35	50.0	1997	1	OTOF_HUMAN	Q9hc10 homo sapien
24	34.5	49.3	1916	1	RIF1_YEAST	P29539 saccharomyc
25	34	48.6	125	1	UR2G_CYPCA	P06580 cyprinus ca
26	34	48.6	139	1	IGF_MYXGL	P22618 myxine glut
27	34	48.6	156	1	ECP3_MOUSE	O35290 mus musculu
28	34	48.6	227	1	PRRA_RAT	P09320 rattus norv
29	34	48.6	235	1	RPSE_CLOAB	P33657 clostridium
30	34	48.6	403	1	SHBG_RAT	P08689 rattus norv
31	34	48.6	433	1	XTMB_BACSU	P39786 bacillus su
32	34	48.6	460	1	CDS1_SCHPO	Q09170 schizosacch
33	34	48.6	493	1	CAMA_CHICK	P05099 gallus gall
34	34	48.6	515	1	DLTA_STRMU	Q53526 streptococc
35	34	48.6	576	1	ACH2_DROME	P17644 drosophila
36	34	48.6	742	1	PURL_BACSU	P12042 bacillus su
37	34	48.6	821	1	TRKB_MOUSE	P15209 mus musculu
38	34	48.6	821	1	TRKB_RAT	Q63604 rattus norv
39	34	48.6	901	1	PHSG_YEAST	P06738 saccharomyc
40	34	48.6	984	1	DPOL_NPVAC	P18131 autographa
41	34	48.6	986	1	DPOL_NPVBM	P41712 bombyx mori
42	34	48.6	1333	1	CC25_CANAL	P43069 candida alb
43	33	47.1	31	1	LPL_BUCRP	Q53017 buchnera ap
44	33	47.1	93	1	YO08_BPHP1	P51709 bacteriopha
45	33	47.1	99	1	PD11_MOUSE	P56983 mus musculu

ALIGNMENTS

RESULT 1
SOML_TETMU
ID SOML_TETMU STANDARD; PRT; 229 AA.
AC Q9I9H4;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Somatolactin precursor (SL).
OS Tetraodon miurus (Congo puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodon.
OX NCBI_TaxID=94908;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA Rand-Weaver M., May D.;
RT "Cloning and sequencing of Tetraodon miurus somatolactin.";
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC ---

CC ---
DR EMBL; AF253066; AAF64522.1; -.
DR InterPro; IPR001400; SOMATOTROPIN.
DR Pfam; PF00103; hormone; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 229 SOMATOLACTIN.
FT DISULFID 26 36 BY SIMILARITY.
FT DISULFID 87 203 BY SIMILARITY.
FT DISULFID 220 228 BY SIMILARITY.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 229 AA; 26125 MW; C10CCF295D28C447 CRC64;

Query Match 57.1%; Score 40; DB 1; Length 229;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LKLLLKLLKC 14
:::|||||||
Db 211 IEILLKLLKC 220

```
RESULT   2
SOM1_SPAAU
ID   SOM1_SPAAU     STANDARD;      PRT;   231 AA.
AC   P54863;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Somatolactin 1 precursor (SL).
OS   Sparus aurata (Gilthead sea bream).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sparidae; Sparus.
OX   NCBI_TaxID=8175;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=97114243; PubMed=8954766;
RA   Astola A., Pendon C., Ortiz M., Valdivia M.M.;
RT   "Cloning and expression of somatolactin, a pituitary hormone related
RT   to growth hormone and prolactin from gilthead seabream, Sparus
RT   aurata.";
RL   Gen. Comp. Endocrinol. 104:330-336(1996).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L49205; AAA98734.1; -.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL       1     24       POTENTIAL.
FT   CHAIN       25    231       SOMATOLACTIN 1.
FT   DISULFID    29     39       BY SIMILARITY.
FT   DISULFID    89    205       BY SIMILARITY.
FT   DISULFID   222    230       BY SIMILARITY.
FT   CARBOHYD   145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   231 AA;  26961 MW;  67A44E7D43E02504 CRC64;


  Query Match             57.1%;  Score 40;  DB 1;  Length 231;
  Best Local Similarity   70.0%;  Pred. No. 12;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 LKLLLKLLKC 14
           :::|||||||
Db     213 MEILLKLLKC 222


RESULT   3
SOM2_SPAAU
ID   SOM2_SPAAU     STANDARD;      PRT;   231 AA.
AC   P79894;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Somatolactin 2 precursor (SL).
OS   Sparus aurata (Gilthead sea bream).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sparidae; Sparus.
OX   NCBI_TaxID=8175;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RA   Cavari B., Funkenstein B., Kawauchi H.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y11144; CAA72031.1; -.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL       1     24       POTENTIAL.
FT   CHAIN       25    231       SOMATOLACTIN 2.
FT   DISULFID    29     39       BY SIMILARITY.
FT   DISULFID    89    205       BY SIMILARITY.
FT   DISULFID   222    230       BY SIMILARITY.
FT   CARBOHYD   145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   231 AA;  26765 MW;  09C774C68DE08BA1 CRC64;


  Query Match             57.1%;  Score 40;  DB 1;  Length 231;
  Best Local Similarity   70.0%;  Pred. No. 12;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 LKLLLKLLKC 14
           :::|||||||
Db     213 MEILLKLLKC 222


RESULT   4
SOML_SCIOC
ID   SOML_SCIOC     STANDARD;      PRT;   231 AA.
AC   Q9YGK7;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Somatolactin precursor (SL).
OS   Sciaenops ocellatus (Red drum).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sciaenidae; Sciaenops.
OX   NCBI_TaxID=76340;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=99102183; PubMed=9882545;
RA   Zhu Y., Yoshiura Y., Kikuchi K., Aida K., Thomas P.;
RT   "Cloning and phylogenetic relationship of red drum somatolactin cDNA
RT   and effects of light on pituitary somatolactin mRNA expression.";
RL   Gen. Comp. Endocrinol. 113:69-79(1999).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF062520; AAD17534.1; -.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    231       SOMATOLACTIN.
FT   DISULFID     29     39       BY SIMILARITY.
FT   DISULFID     89    205       BY SIMILARITY.
FT   DISULFID    222    230       BY SIMILARITY.
FT   CARBOHYD    145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   231 AA;  26658 MW;  4FB039DEE6EDBD01 CRC64;


  Query Match             57.1%;  Score 40;  DB 1;  Length 231;
  Best Local Similarity   70.0%;  Pred. No. 12;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 LKLLLKLLKC 14
           :::|||||||
Db     213 MEILLKLLKC 222


RESULT   5
SOML_SIGGU
ID   SOML_SIGGU     STANDARD;      PRT;   231 AA.
AC   Q9PWG4;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Somatolactin precursor (SL).
OS   Siganus guttatus (Rabbitfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
OC   Acanthuroidei; Siganidae; Siganus.
OX   NCBI_TaxID=92439;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RA   Ayson F.G., de Jesus E.T., Amemiya Y., Moriyama S., Hirano T.,
RA   Kawauchi H.;
RT   "Isolation and cDNA cloning of somatolactin in rabbitfish (Siganus
RT   guttatus).";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AB026186; BAA83467.1; -.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    231       SOMATOLACTIN.
FT   DISULFID     29     39       BY SIMILARITY.
FT   DISULFID     89    205       BY SIMILARITY.
FT   DISULFID    222    230       BY SIMILARITY.
FT   CARBOHYD     35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   231 AA;  26594 MW;  09A9CD5EE13840AC CRC64;


  Query Match             57.1%;  Score 40;  DB 1;  Length 231;
  Best Local Similarity   70.0%;  Pred. No. 12;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 LKLLLKLLKC 14
           :::|||||||
Db     213 MEILLKLLKC 222


RESULT   6
SOML_CARAU
ID   SOML_CARAU     STANDARD;      PRT;   230 AA.
AC   P79697;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Somatolactin precursor (SL).
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97242175; PubMed=9125164;
RA   Cheng K.W., Chan Y.H., Chen Y.D., Yu K.L., Chan K.M.;
RT   "Sequence of a cDNA clone encoding a novel somatolactin in goldfish,
RT   Carassius auratus.";
RL   Biochem. Biophys. Res. Commun. 232:282-287(1997).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U72940; AAC60098.1; -.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PROSITE; PS00266; SOMATOTROPIN_1; FALSE_NEG.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    230       SOMATOLACTIN.
FT   DISULFID     28     38       BY SIMILARITY.
FT   DISULFID     87    202       BY SIMILARITY.
FT   DISULFID    219    227       BY SIMILARITY.
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   230 AA;  25735 MW;  CBC5DB347C6116DC CRC64;


  Query Match             52.9%;  Score 37;  DB 1;  Length 230;
  Best Local Similarity   70.0%;  Pred. No. 36;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 LKLLLKLLKC 14
           :: |||||||
Db     210 IQTLLKLLKC 219
```

```
RESULT   7
Y035_TREPA
ID   Y035_TREPA      STANDARD;      PRT;   238 AA.
AC   O83078;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable metal transport system ATP-binding protein TP0035.
GN   TP0035.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NICHOLS;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   -!- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM
CC       TP0034/TP0035/TP0036 FOR A METAL. PROBABLY RESPONSIBLE FOR ENERGY
CC       COUPLING TO THE TRANSPORT SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE001188; AAC65030.1; -.
DR   TIGR; TP0035; -.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; FALSE_NEG.
KW   Hypothetical protein; Transport; Inner membrane; ATP-binding;
KW   Complete proteome.
FT   NP_BIND      44     51       ATP (POTENTIAL).
SQ   SEQUENCE   238 AA;  26460 MW;  673E7B4882BE4D29 CRC64;

  Query Match             52.9%;  Score 37;  DB 1;  Length 238;
  Best Local Similarity   80.0%;  Pred. No. 37;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 LLKLLLKLLK 13
           |:||:|||||
Db      53 LVKLVLKLLK 62


RESULT   8
LP61_EIMTE
ID   LP61_EIMTE      STANDARD;      PRT;   255 AA.
AC   P15714;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Antigen LPMC-61 (Fragment).
OS   Eimeria tenella.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
OC   Eimeria.
OX   NCBI_TaxID=5802;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Sporozoite;
RX   MEDLINE=90348718; PubMed=2200963;
RA   Ko C., Smith C.K. II, McDonell M.;
RT   "Identification and characterization of a target antigen of a
RT   monoclonal antibody directed against Eimeria tenella merozoites.";
RL   Mol. Biochem. Parasitol. 41:53-64(1990).
CC   -!- FUNCTION: UNKNOWN. THE GLN-RICH TANDEM REPEATS MAY BE IMPORTANT
CC       FOR AN UNKNOWN ASPECT OF THE PARASITIC LIFE CYCLE. MAY BE AN
CC       IMPORTANT IMMUNOGEN.
CC   -!- SUBUNIT: MAY BE COVALENTLY LINKED BY DISULFIDE BONDS TO OTHER
CC       POLYPEPTIDES TO FORM THE 80 kDa ANTIGEN.
CC   -!- DEVELOPMENTAL STAGE: PRESENT IN ALL STAGES THROUGHOUT THE
CC       SPORULATION OF THE OOCYSTS AND IN THE SPOROZOITES FOLLOWING
CC       EXCYSTATION.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M30933; AAA29079.1; -.
DR   PIR; A60637; A60637.
KW   Antigen; Sporozoite; Repeat; Sporulation.
FT   NON_TER       1      1
FT   DOMAIN       18    210       12 X APPROXIMATE TANDEM REPEATS, GLN-
FT                                 RICH.
FT   REPEAT       18     48       1.
FT   REPEAT       49     57       2.
FT   REPEAT       58     65       3.
FT   REPEAT       66     78       4.
FT   REPEAT       79     90       5.
FT   REPEAT       91    103       6.
FT   REPEAT      104    140       7.
FT   REPEAT      141    152       8.
FT   REPEAT      153    164       9.
FT   REPEAT      165    172       10.
FT   REPEAT      173    192       11.
FT   REPEAT      193    210       12.
FT   NON_TER     255    255
SQ   SEQUENCE   255 AA;  31267 MW;  8C5E6005FFFC2DB3 CRC64;

  Query Match             52.9%;  Score 37;  DB 1;  Length 255;
  Best Local Similarity   81.8%;  Pred. No. 39;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLL 12
           :|||||||| ||
Db       2 RLLLKLLLLLL 12


RESULT   9
COPG_YEAST
ID   COPG_YEAST      STANDARD;      PRT;   935 AA.
AC   P32074;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Coatomer gamma subunit (Gamma-coat protein) (Gamma-COP).
GN   SEC21 OR YNL287W OR N0543.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=93096049; PubMed=1461285;
RA   Hosobuchi M.M., Kreis T., Schekman R.;
RT   "SEC21 is a gene required for ER to Golgi protein transport that
RT   encodes a subunit of a yeast coatomer.";
RL   Nature 360:603-605(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Messenguy F., Dubois E., Vierendeels F., Scherens B., Pierard A.,
RA   Glansdorff N.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE COATOMER IS A CYTOSOLIC PROTEIN COMPLEX THAT BINDS
CC       TO DILYSINE MOTIFS AND REVERSIBLY ASSOCIATES WITH GOLGI NON-
CC       CLATHRIN-COATED VESICLES, WHICH FURTHER MEDIATE BIOSYNTHETIC
CC       PROTEIN TRANSPORT FROM THE ER, VIA THE GOLGI UP TO THE TRANS GOLGI
CC       NETWORK. COATOMER COMPLEX IS REQUIRED FOR BUDDING FROM GOLGI
CC       MEMBRANES, AND IS ESSENTIAL FOR THE RETROGRADE GOLGI-TO-ER
CC       TRANSPORT OF DILYSINE-TAGGED PROTEINS (BY SIMILARITY).
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX THAT CONSISTS OF AT LEAST THE ALPHA,
CC       BETA, BETA', GAMMA, DELTA, EPSILON AND ZETA SUBUNITS.
CC   -!- SUBCELLULAR LOCATION: THE COATOMER IS CYTOPLASMIC OR POLYMERIZED
CC       ON THE CYTOPLASMIC SIDE OF THE GOLGI, AS WELL AS ON THE
CC       VESICLES/BUDS ORIGINATING FROM IT (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE COPG FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M59708; AAA34598.1; -.
DR   EMBL; Z71563; CAA96204.1; -.
DR   PIR; A33151; A33151.
DR   PIR; S28915; S28915.
DR   SGD; S0005231; SEC21.
DR   InterPro; IPR002553; Adaptin_N.
DR   Pfam; PF01602; Adaptin_N; 1.
KW   Transport; Protein transport; Golgi stack; Membrane.
FT   CONFLICT    353    353       D -> N (IN REF. 1).
SQ   SEQUENCE   935 AA;  104830 MW;  99DC7D737D4EF761 CRC64;


  Query Match             52.9%;  Score 37;  DB 1;  Length 935;
  Best Local Similarity   58.3%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLL 12
           |:||:  ||:||
Db      41 CRLLISRLLRLL 52


RESULT  10
EAG_BACSU
ID   EAG_BACSU      STANDARD; .     PRT;   143 AA.
AC   P06630;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 16.4 kDa protein in SPO0E 3'region.
GN   EAG.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=88260878; PubMed=2838724;
RA   Perego M., Hoch J.A.;
RT   "Isolation and sequence of the spo0E gene: its role in initiation of
RT   sporulation in Bacillus subtilis.";
RL   Mol. Microbiol. 1:125-132(1987).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y00526; CAA68584.1; -.
DR   EMBL; Z99111; CAB13238.1; -.
DR   PIR; S03747; S03747.
DR   SubtiList; BG10770; eag.
KW   Hypothetical protein; Sporulation; Complete proteome.
SQ   SEQUENCE   143 AA;  16429 MW;  D7410B50963D7A75 CRC64;


  Query Match             51.4%;  Score 36;  DB 1;  Length 143;
  Best Local Similarity   63.6%;  Pred. No. 34;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 LLLKLLLKLLK 13
           ||||:||: :|
Db     114 LLLKMLLRFIK 124


RESULT  11
SOML_HIPHI
ID   SOML_HIPHI     STANDARD;      PRT;   230 AA.
AC   P45641;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Somatolactin precursor (SL).
OS   Hippoglossus hippoglossus (Atlantic halibut).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC   Pleuronectoidei; Pleuronectidae; Hippoglossus.
OX   NCBI_TaxID=8267;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93372995; PubMed=7689905;
RA   Iraqi F., Gong Z., Hew C.L., Crim L.W.;
RT   "Isolation and characterization of somatolactin genes from two cold
RT   water marine teleosts, lumpfish (Cyclopterus lumpus) and halibut
RT   (Hippoglossus hippoglossus).";
RL   Mol. Mar. Biol. Biotechnol. 2:96-103(1993).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L02117; AAC38003.1; -.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    230       SOMATOLACTIN.
FT   DISULFID     28     38       BY SIMILARITY.
FT   DISULFID     88    204       BY SIMILARITY.
FT   DISULFID    221    229       BY SIMILARITY.
```

```
FT   CARBOHYD    137    137       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   230 AA;  26519 MW;  1E3C16622F75946A CRC64;


  Query Match             51.4%;  Score 36;  DB 1;  Length 230;
  Best Local Similarity   60.0%;  Pred. No. 51;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 LKLLLKLLKC 14
          ::: ||||||
Db    212 MEIFLKLLKC 221


RESULT  12
SOML_SOLSE
ID   SOML_SOLSE     STANDARD;      PRT;   230 AA.
AC   P45642;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Somatolactin precursor (SL).
OS   Solea senegalensis (Sole).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC   Soleoidei; Soleidae; Solea.
OX   NCBI_TaxID=28829;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=95011619; PubMed=7926805;
RA   Pendon C., Martinez-Barbera J.P., Valdivia M.M.;
RT   "Cloning of a somatolactin-encoding cDNA from sole (Solea
RT   senegalensis).";
RL   Gene 147:227-230(1994).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U06753; AAA61873.1; -.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    230       SOMATOLACTIN.
FT   DISULFID     28     38       BY SIMILARITY.
FT   DISULFID     88    204       BY SIMILARITY.
FT   DISULFID    221    229       BY SIMILARITY.
FT   CARBOHYD     34     34       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   230 AA;  26586 MW;  ACDB39117EE3D49A CRC64;


  Query Match             51.4%;  Score 36;  DB 1;  Length 230;
  Best Local Similarity   60.0%;  Pred. No. 51;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 LKLLLKLLKC 14
          ::: ||||||
Db    212 MEIFLKLLKC 221


RESULT  13
SOML_PAROL
ID   SOML_PAROL     STANDARD;      PRT;   231 AA.
AC   P20362;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Somatolactin precursor (SL).
OS   Paralichthys olivaceus (Flounder).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC   Pleuronectoidei; Paralichthyidae; Paralichthys.
OX   NCBI_TaxID=8255;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 25-47.
RC   TISSUE=Pituitary;
RX   MEDLINE=90272707; PubMed=2349240;
RA   Ono M., Takayama Y., Rand-Weaver M., Sakata S., Yasunaga T., Noso T.,
RA   Kawauchi H.;
RT   "cDNA cloning of somatolactin, a pituitary protein related to growth
RT   hormone and prolactin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:4330-4334(1990).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PITUITARY GLAND.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M33696; AAA49445.1; -.
DR   EMBL; M33695; AAA49444.1; -.
DR   PIR; A35793; A35793.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    231       SOMATOLACTIN.
FT   DISULFID     29     39       BY SIMILARITY.
FT   DISULFID     89    205       BY SIMILARITY.
FT   DISULFID    222    230       BY SIMILARITY.
FT   CARBOHYD    145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   231 AA;  26731 MW;  8C7EDAA6BB912BAB CRC64;


  Query Match             51.4%;  Score 36;  DB 1;  Length 231;
  Best Local Similarity   60.0%;  Pred. No. 52;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 LKLLLKLLKC 14
          ::: ||||||
Db    213 MEIFLKLLKC 222


RESULT  14
PRR1_BOVIN
ID   PRR1_BOVIN     STANDARD;      PRT;   238 AA.
AC   P05402;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Placental prolactin-related protein I precursor (PRC-I).
```

```
GN   PRP1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87289662; PubMed=3475696;
RA   Schuler L.A., Hurley W.L.;
RT   "Molecular cloning of a prolactin-related mRNA expressed in bovine
RT   placenta.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:5650-5654(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89251077; PubMed=2721368;
RA   Ebbitt D.M., Hurley W.L., Kessler M.A., McDonald D.J., Schuler L.A.;
RT   "Characterization of the gene corresponding to bovine placental
RT   prolactin-related cDNA I: evolutionary implications.";
RL   DNA 8:161-169(1989).
CC   -!- FUNCTION: PLACENTAL PROLACTIN-RELATED PROTEINS MAY PLAY A
CC       SPECIFIC ROLE DURING GESTATION.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02944; AAA30726.1; -.
DR   EMBL; M25494; AAA30727.1; -.
DR   EMBL; M25491; AAA30727.1; JOINED.
DR   EMBL; M25492; AAA30727.1; JOINED.
DR   EMBL; M25493; AAA30727.1; JOINED.
DR   PIR; A31417; A31417.
DR   HSSP; Q28632; 1AN3.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Placenta; Signal.
FT   SIGNAL        1     36       POTENTIAL.
FT   CHAIN        37    238       PLACENTAL PROLACTIN-RELATED PROTEIN I.
FT   DISULFID     97    215       BY SIMILARITY.
FT   DISULFID    232    238       BY SIMILARITY.
FT   CONFLICT    201    201       A -> D (IN REF. 2).
SQ   SEQUENCE   238 AA;  27675 MW;  EC3609F025BEF808 CRC64;


  Query Match             51.4%;  Score 36;  DB 1;  Length 238;
  Best Local Similarity   57.1%;  Pred. No. 53;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLLKC 14
           | ||| |:  || |
Db      21 CLLLLLLMSNLLLC 34


RESULT  15
Y147_MYCGE
ID   Y147_MYCGE     STANDARD;      PRT;   375 AA.
AC   P47393;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein MG147.
GN   MG147.
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RX   MEDLINE=96026346; PubMed=7569993;
RA   Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA   Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA   Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA   Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA   Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA   Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT   "The minimal gene complement of Mycoplasma genitalium.";
RL   Science 270:397-403(1995).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U39695; AAC71365.1; -.
DR   TIGR; MG147; -.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM     21     41       POTENTIAL.
FT   TRANSMEM     66     86       POTENTIAL.
FT   TRANSMEM    161    181       POTENTIAL.
FT   TRANSMEM    203    223       POTENTIAL.
FT   TRANSMEM    234    254       POTENTIAL.
FT   TRANSMEM    289    309       POTENTIAL.
FT   TRANSMEM    338    358       POTENTIAL.
SQ   SEQUENCE   375 AA;  43188 MW;  A14AF07D574E8046 CRC64;


  Query Match             51.4%;  Score 36;  DB 1;  Length 375;
  Best Local Similarity   66.7%;  Pred. No. 80;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLL 12
           || || ||: ||
Db      17 CKALLLLLIPLL 28


Search completed: June 17, 2002, 12:44:48
Job time: 303 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:44:21 ; Search time 73.61 Seconds
(without alignments)
32.902 Million cell updates/sec

Title: US-09-367-714A-92
Perfect score: 70
Sequence: 1 CKLLLKLLLKLLKC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42	60.0	1300	5	Q9NKD6	Q9nkd6 drosophila
2	42	60.0	1360	5	Q9NDI1	Q9ndi1 drosophila
3	41	58.6	295	9	Q94M45	Q94m45 streptococc
4	40	57.1	106	5	Q9VUG7	Q9vug7 drosophila
5	40	57.1	169	16	Q9Z8R1	Q9z8r1 chlamydia p
6	40	57.1	207	13	Q9PSN4	Q9psn4 sparus aura
7	40	57.1	216	13	Q9DET0	Q9det0 dicentrarch
8	40	57.1	314	3	Q05867	Q05867 saccharomyc
9	40	57.1	756	11	Q9CUE5	Q9cue5 mus musculu
10	40	57.1	988	5	Q9N8K2	Q9n8k2 trypanosoma
11	40	57.1	1022	10	Q9FL22	Q9fl22 arabidopsis
12	40	57.1	1880	5	O18465	O18465 hirudo medi
13	40	57.1	2481	10	Q9FR53	Q9fr53 arabidopsis
14	40	57.1	2513	10	Q9LPM4	Q9lpm4 arabidopsis
15	39	55.7	641	10	Q9ZT73	Q9zt73 arabidopsis
16	38	54.3	219	13	Q99370	Q99370 gallus gall
17	38	54.3	248	11	Q9ESC4	Q9esc4 mus musculu
18	38	54.3	331	11	Q9ESC5	Q9esc5 mus musculu
19	38	54.3	437	11	Q9CZU4	Q9czu4 mus musculu
20	38	54.3	437	11	Q925U1	Q925u1 mus musculu
21	38	54.3	513	8	Q9MV03	Q9mv03 pharus lati
22	38	54.3	518	10	Q94D22	Q94d22 oryza sativ
23	38	54.3	754	10	Q9ZQZ2	Q9zqz2 arabidopsis
24	38	54.3	1119	5	Q19190	Q19190 caenorhabdi
25	38	54.3	2045	5	Q9W444	Q9w444 drosophila
26	38	54.3	2162	12	O91940	O91940 bovine resp
27	38	54.3	2162	12	Q9WKU5	Q9wku5 bovine resp
28	37	52.9	83	16	Q928W1	Q928w1 listeria in
29	37	52.9	96	10	Q9LWS8	Q9lws8 oryza sativ
30	37	52.9	116	3	O74917	O74917 schizosacch
31	37	52.9	137	16	Q97MS6	Q97ms6 clostridium
32	37	52.9	161	10	Q9LZ38	Q9lz38 arabidopsis
33	37	52.9	191	2	Q9R9K5	Q9r9k5 paracoccus
34	37	52.9	275	16	Q9ZDI6	Q9zdi6 rickettsia
35	37	52.9	398	5	Q22902	Q22902 caenorhabdi
36	37	52.9	481	5	O45540	O45540 caenorhabdi
37	37	52.9	515	8	Q9GHC4	Q9ghc4 paris thibe
38	37	52.9	1030	5	Q19645	Q19645 caenorhabdi
39	37	52.9	1252	5	Q9Y0D0	Q9y0d0 hydra atten
40	36.5	52.1	846	2	Q9AIP5	Q9aip5 candidatus
41	36	51.4	36	10	Q9M3S5	Q9m3s5 arabidopsis
42	36	51.4	53	8	Q9BC70	Q9bc70 sargassum p
43	36	51.4	74	10	Q94LQ5	Q94lq5 oryza sativ
44	36	51.4	120	3	Q05715	Q05715 saccharomyc
45	36	51.4	143	4	O15412	O15412 homo sapien

ALIGNMENTS

RESULT 1
Q9NKD6
ID Q9NKD6 PRELIMINARY; PRT; 1300 AA.
AC Q9NKD6; Q9VJU3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 144.0 KDA PROTEIN (RK GENE PRODUCT).
GN RK OR BG:DS00180.13 OR CG8930.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, AND CN BW SP;
RX MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Celniker S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, AND CN BW SP;
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003408; AAF44846.1; -.
DR   EMBL; AE003642; AAF53367.2; -.
DR   FlyBase; FBgn0003255; rk.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00370; LRR; 1.
DR   SMART; SM00369; LRR_TYP; 4.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   1300 AA;  144031 MW;  B4B9E39F942FA0B3 CRC64;


  Query Match             60.0%;  Score 42;  DB 5;  Length 1300;
  Best Local Similarity   75.0%;  Pred. No. 85;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLL 12
           | |||:|||:|| 
Db      14 CPLLLQLLLQLL 25


RESULT   2
Q9NDI1
ID   Q9NDI1      PRELIMINARY;      PRT;  1360 AA.
AC   Q9NDI1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLYCOPROTEIN HORMONE RECEPTOR II.
GN   RK OR BG:DS00180.13 OR CG8930.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON S.; TISSUE=WHOLE ANIMAL;
RX   MEDLINE=20359836; PubMed=10899142;
RA   Eriksen K.K., Hauser F., Schiott M., Pedersen K.-M., Soendergaard L.,
RA   Grimmelikhuijzen C.J.P.;
RT   "Molecular Cloning, Genomic Organization, Developmental Regulation,
RT   and a Knock-Out Mutant of a Novel Leu-Rich Repeats-Containing G
RT   Protein-Coupled Receptor (DLGR-2) from Drosophila melanogaster.";
RL   Genome Res. 10:924-938(2000).
DR   EMBL; AF142343; AAF66608.1; -.
DR   HSSP; Q57815; 1D3Y.
DR   FlyBase; FBgn0003255; rk.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 14.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00370; LRR; 2.
DR   SMART; SM00369; LRR_TYP; 5.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   1360 AA;  150731 MW;  7D435155B4F6F612 CRC64;


  Query Match             60.0%;  Score 42;  DB 5;  Length 1360;
  Best Local Similarity   75.0%;  Pred. No. 88;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLL 12
           | |||:|||:|| 
Db      14 CPLLLQLLLQLL 25


RESULT   3
Q94M45
ID   Q94M45      PRELIMINARY;      PRT;  295 AA.
AC   Q94M45;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 32.7 KDA PROTEIN.
OS   Streptococcus pneumoniae bacteriophage MM1.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX   NCBI_TaxID=120574;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Obregon V., Garcia P., Garcia E., Lopez R., Garcia J.L.;
RT   "Complete nucleotide sequence and analysis of the temperate
RT   bacteriophage MM1 genome of Streptococcus pneumoniae.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ302074; CAC48100.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   295 AA;  32657 MW;  71A01307E78B6ACF CRC64;


  Query Match             58.6%;  Score 41;  DB 9;  Length 295;
  Best Local Similarity   66.7%;  Pred. No. 35;
  Matches    8;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLL 12
           ||||:::| || 
Db     267 CKLLLRVVLTLL 278
```

```
RESULT   4
Q9VUG7
ID   Q9VUG7       PRELIMINARY;      PRT;   106 AA.
AC   Q9VUG7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CG13476 PROTEIN.
GN   CG13476.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003533; AAF49716.2; -.
DR   FlyBase; FBgn0036441; CG13476.
SQ   SEQUENCE   106 AA;  12193 MW;  92041B063C951FC4 CRC64;


  Query Match             57.1%;  Score 40;  DB 5;  Length 106;
  Best Local Similarity   46.2%;  Pred. No. 22;
  Matches    6;  Conservative    7;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLLK 13
           |:|||::|:|::
Db      77 CELLLEILIKIVE 89


RESULT   5
Q9Z8R1
ID   Q9Z8R1       PRELIMINARY;      PRT;   169 AA.
AC   Q9Z8R1;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 18.2 KDA PROTEIN.
GN   CPN0277 OR CPJ0277 OR CP0481.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
DR   EMBL; AE001613; AAD18426.1; -.
DR   EMBL; AE002208; AAF38311.1; -.
DR   EMBL; AP002546; BAA98487.1; -.
DR   TIGR; CP0481; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   169 AA;  18203 MW;  4A3B2967C18A7424 CRC64;


  Query Match             57.1%;  Score 40;  DB 16;  Length 169;
  Best Local Similarity   50.0%;  Pred. No. 32;
  Matches    6;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLLKLLLKLLKC 14
           |:|::::||:||
Db      94 LVLQVIIKLIKC 105


RESULT   6
Q9PSN4
ID   Q9PSN4       PRELIMINARY;      PRT;   207 AA.
AC   Q9PSN4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SOMATOLACTIN, SL.
OS   Sparus aurata (Gilthead sea bream).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sparidae; Sparus.
OX   NCBI_TaxID=8175;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95291367; PubMed=7773329;
RA   Cavari B., Noso T., Kawauchi H.;
```

```
RT   "Somatolactin, a novel pituitary protein: isolation and
RT   characterization from Sparus aurata.";
RL   Mol. Mar. Biol. Biotechnol. 4:117-122(1995).
DR   HSSP; P01241; 1AXI.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
SQ   SEQUENCE   207 AA;  23888 MW;  FDA8BBBEC9737271 CRC64;


  Query Match             57.1%;  Score 40;  DB 13;  Length 207;
  Best Local Similarity   70.0%;  Pred. No. 38;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 LKLLLKLLKC 14
           :::|||||||
Db     189 MEILLKLLKC 198


RESULT   7
Q9DET0
ID   Q9DET0      PRELIMINARY;      PRT;   216 AA.
AC   Q9DET0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SOMATOLACTIN PRECURSOR (FRAGMENT).
GN   SL.
OS   Dicentrarchus labrax (European sea bass).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Moronidae; Dicentrarchus.
OX   NCBI_TaxID=13489;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Company R., Calduch-Giner J.A., Mingarro M., Perez-Sanchez J.;
RT   "cDNA cloning and sequence of European sea bass (Dicentrarchus labrax)
RT   somatolactin.";
RL   Comp. Biochem. Physiol. 127:183-192(2000).
DR   EMBL; AJ277390; CAC16116.1; -.
DR   HSSP; P01241; 1AXI.
DR   InterPro; IPR001400; SOMATOTROPIN.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL        <1     9       POTENTIAL.
FT   CHAIN         10     216     SOMATOLACTIN.
SQ   SEQUENCE   216 AA;  25010 MW;  95CBB324A6069F00 CRC64;


  Query Match             57.1%;  Score 40;  DB 13;  Length 216;
  Best Local Similarity   70.0%;  Pred. No. 39;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 LKLLLKLLKC 14
           :::|||||||
Db     198 MEILLKLLKC 207


RESULT   8
Q05867
ID   Q05867      PRELIMINARY;      PRT;   314 AA.
AC   Q05867;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CHROMOSOME XII COSMID 8003.
GN   YLR283W OR L8003.9.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97313267; PubMed=9169871;
RA   Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
RA   Benes V., Bruckner M., Delius H., Dubois E., Dusterhoft A.,
RA   Entian K.D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA   Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Kotter P.,
RA   Louis E.J., Messenguy F., Mewes H.W., Miosga T., Mostl D.,
RA   Muller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA   Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA   Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA   Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA   Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA   Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL   Nature 387:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Pauley A.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Waterston R.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Cherry J.M.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U17243; AAB67328.1; -.
DR   SGD; S0004273; YLR283W.
SQ   SEQUENCE   314 AA;  36574 MW;  2CEFDB5C73D57B08 CRC64;


  Query Match             57.1%;  Score 40;  DB 3;  Length 314;
  Best Local Similarity   58.3%;  Pred. No. 54;
  Matches    7;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLL 12
           ||::: |||:||
Db     122 CKIIIALLLQLL 133


RESULT   9
Q9CUE5
ID   Q9CUE5      PRELIMINARY;      PRT;   756 AA.
AC   Q9CUE5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   4931427F14RIK PROTEIN (FRAGMENT).
GN   4931427F14RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=TESTIS;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
```

```
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK016477; BAB30259.1; -.
DR   MGD; MGI:1921612; 4931427F14Rik.
FT   NON_TER     756    756
SQ   SEQUENCE   756 AA;  86030 MW;  19A6BE7FD7853652 CRC64;


  Query Match             57.1%;  Score 40;  DB 11;  Length 756;
  Best Local Similarity   69.2%;  Pred. No. 1.1e+02;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLKC 14
           || | |: |||||
Db     118 KLTLPLIGKLLKC 130


RESULT  10
Q9N8K2
ID   Q9N8K2      PRELIMINARY;      PRT;   988 AA.
AC   Q9N8K2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 107.3 KDA PROTEIN.
GN   CHR1.297.
OS   Trypanosoma brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5691;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TREU927;
RA   Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
RA   Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
RA   Gerrard C., Rajandream M.A., Barrell B.G.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL359782; CAB95541.1; -.
DR   InterPro; IPR000847; HTH_LysR.
DR   PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
KW   Hypothetical protein.
SQ   SEQUENCE   988 AA;  107318 MW;  EFB3A38C56AE6E85 CRC64;


  Query Match             57.1%;  Score 40;  DB 5;  Length 988;
  Best Local Similarity   61.5%;  Pred. No. 1.4e+02;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLKC 14
           :| |||:|:||:|
Db     432 ELHLKLVLRLLQC 444


RESULT  11
Q9FL22
ID   Q9FL22      PRELIMINARY;      PRT;  1022 AA.
AC   Q9FL22;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   GENOMIC DNA, CHROMOSOME 5, P1 CLONE:MPL12.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=98344145; PubMed=9679202;
RA   Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT   features of the regions of 1,381,565 bp covered by twenty one
RT   physically assigned P1 and TAC clones.";
RL   DNA Res. 5:131-145(1998).
DR   EMBL; AB010698; BAB11095.1; -.
DR   InterPro; IPR003107; HAT.
DR   SMART; SM00386; HAT; 4.
SQ   SEQUENCE   1022 AA;  117365 MW;  B59AC43225E4A17F CRC64;


  Query Match             57.1%;  Score 40;  DB 10;  Length 1022;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+02;
  Matches    8;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CKLLLKLLLKLL 12
           |||||: |::||
Db     477 CKLLLEELMRLL 488


RESULT  12
O18465
ID   O18465      PRELIMINARY;      PRT;  1880 AA.
AC   O18465;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TRACTIN.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92198663; PubMed=1550678;
RA   Johansen K.M., Kopp D.M., Jellies J., Johansen J.;
RT   "Tract formation and axon fasciculation of molecularly distinct
RT   peripheral neuron subpopulations during leech embryogenesis.";
RL   Neuron 8:559-572(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97362067; PubMed=9214388;
RA   Huang Y., Jellies J., Johansen K.M., Johansen J.;
RT   "Differential glycosylation of tractin and LeechCAM, two novel Ig
RT   superfamily members, regulates neurite extension and fascicle
RT   formation.";
RL   J. Cell Biol. 138:143-157(1997).
DR   EMBL; U92813; AAC47654.1; -.
DR   HSSP; P20241; 1CFB.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR003962; FnIII_repeat.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00041; fn3; 4.
DR   Pfam; PF00047; ig; 6.
DR   PRINTS; PR00014; FNTYPEIII.
DR   SMART; SM00060; FN3; 4.
DR   SMART; SM00408; IGc2; 4.
DR   SMART; SM00410; IG_like; 2.
KW   Immunoglobulin domain; Repeat.
SQ   SEQUENCE   1880 AA;  199865 MW;  174EC84DAC540DF0 CRC64;


  Query Match             57.1%;  Score 40;  DB 5;  Length 1880;
  Best Local Similarity   64.3%;  Pred. No. 2.4e+02;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy       1 CKLLLKLLLKLLKC 14
           | ||| ||| :| |
Db    1758 CLLLLLLLLLILLC 1771
```

```
RESULT  13
Q9FR53
ID   Q9FR53       PRELIMINARY;      PRT;  2481 AA.
AC   Q9FR53;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PTOR.
GN   TOR.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Menand B., Nussaume L., Meyer C., Desnos T., Bouchez D., Robaglia C.;
RT   "Mutation in AtTOR affects embryo development.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF178967; AAG43423.1; -.
DR   HSSP; P42345; 1FAP.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
SQ   SEQUENCE   2481 AA;  279187 MW;  DA663EA9A9366F93 CRC64;
```

```
  Query Match             57.1%;  Score 40;  DB 10;  Length 2481;
  Best Local Similarity   90.9%;  Pred. No. 3.1e+02;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 LLLKLLLKLLK 13
           ||| |||||||
Db     785 LLLGLLLKLLK 795
```

```
RESULT  14
Q9LPM4
ID   Q9LPM4       PRELIMINARY;      PRT;  2513 AA.
AC   Q9LPM4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   F2J10.9 PROTEIN.
GN   F2J10.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Sakano H., Liu S.X., Yu G., Lee J., Lenz C., Pham P., Toriumi M.,
RA   Chin C., Chiou J., Choi E., Chung M., Gonzalez A., Howng B., Liu A.,
RA   Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA   Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA   Miranda M., Nguyen M., Palm C., Shinn P., Southwick A., Davis R.W.,
RA   Ecker J.R., Federspiel N.A., Theologis A.;
RT   "The sequence of BAC F2J10 from Arabidopsis thaliana chromosome 1.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC015445; AAF76442.1; -.
DR   HSSP; P42345; 1FAP.
DR   InterPro; IPR003151; FAT.
DR   InterPro; IPR003152; FATC.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF02259; FAT; 1.
DR   Pfam; PF02260; FATC; 1.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
SQ   SEQUENCE   2513 AA;  282911 MW;  A4B9740321AC5261 CRC64;
```

```
  Query Match             57.1%;  Score 40;  DB 10;  Length 2513;
  Best Local Similarity   90.9%;  Pred. No. 3.1e+02;
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 LLLKLLLKLLK 13
           ||| |||||||
Db     797 LLLGLLLKLLK 807
```

```
RESULT  15
Q9ZT73
ID   Q9ZT73       PRELIMINARY;      PRT;  641 AA.
AC   Q9ZT73;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 70.5 KDA PROTEIN.
GN   F9H3.7 OR AT4G03450.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Huang E.N., Parnell L.D., de la Bastide M., Schutz K., Habermann K.,
RA   Dedhia N.N., McCombie W.R.;
RT   "Genomic sequence of Arabidopsis Thaliana BAC F9H3, chromosome IV,
RT   18.8 cM.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF071527; AAD11587.1; -.
DR   EMBL; AL161496; CAB77830.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 9.
DR   SMART; SM00248; ANK; 2.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Hypothetical protein; Repeat.
SQ   SEQUENCE   641 AA;  70466 MW;  ECD57973640403CD CRC64;
```

```
  Query Match             55.7%;  Score 39;  DB 10;  Length 641;
  Best Local Similarity   70.0%;  Pred. No. 1.4e+02;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 LKLLLKLLKC 14
           :|: ||||||
Db     304 VKIFLKLLKC 313
```

Search completed: June 17, 2002, 12:44:23
Job time: 298 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 17, 2002, 12:42:05 ; Search time 34.71 Seconds
(without alignments)
9.852 Million cell updates/sec

Title: US-09-367-714A-92
Perfect score: 70
Sequence: 1 CKLLLKLLLKLLKC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	64.3	20	4	US-09-000-692-1	Sequence 1, Appli
2	40	57.1	14	2	US-08-569-188-8	Sequence 8, Appli
3	40	57.1	14	5	PCT-US94-07019-8	Sequence 8, Appli
4	40	57.1	16	2	US-08-569-188-1	Sequence 1, Appli
5	40	57.1	16	2	US-08-569-188-10	Sequence 10, Appl
6	40	57.1	16	2	US-08-569-188-11	Sequence 11, Appl
7	40	57.1	16	2	US-08-569-188-12	Sequence 12, Appl
8	40	57.1	16	5	PCT-US94-07019-1	Sequence 1, Appli
9	40	57.1	16	5	PCT-US94-07019-10	Sequence 10, Appl
10	40	57.1	16	5	PCT-US94-07019-11	Sequence 11, Appl
11	40	57.1	16	5	PCT-US94-07019-12	Sequence 12, Appl
12	40	57.1	17	2	US-08-569-188-3	Sequence 3, Appli
13	40	57.1	17	2	US-08-818-253-39	Sequence 39, Appl
14	40	57.1	17	4	US-08-818-252-39	Sequence 39, Appl
15	40	57.1	17	5	PCT-US94-07019-3	Sequence 3, Appli
16	40	57.1	18	2	US-08-569-188-5	Sequence 5, Appli
17	40	57.1	18	5	PCT-US94-07019-5	Sequence 5, Appli
18	40	57.1	22	1	US-07-725-331-60	Sequence 60, Appl
19	40	57.1	22	5	PCT-US91-05047-60	Sequence 60, Appl
20	40	57.1	23	2	US-08-290-853-19	Sequence 19, Appl
21	40	57.1	26	1	US-07-725-331-61	Sequence 61, Appl
22	40	57.1	26	5	PCT-US91-05047-61	Sequence 61, Appl
23	40	57.1	30	1	US-07-725-331-62	Sequence 62, Appl
24	40	57.1	30	5	PCT-US91-05047-62	Sequence 62, Appl
25	40	57.1	36	1	US-07-725-331-63	Sequence 63, Appl
26	40	57.1	36	5	PCT-US91-05047-63	Sequence 63, Appl
27	40	57.1	40	2	US-08-687-551-6	Sequence 6, Appli
28	38	54.3	21	1	US-08-944-133-13	Sequence 13, Appl
29	37	52.9	14	1	US-07-725-331-1	Sequence 1, Appli
30	37	52.9	14	5	PCT-US91-05047-1	Sequence 1, Appli
31	37	52.9	16	2	US-08-569-188-2	Sequence 2, Appli
32	37	52.9	16	2	US-08-569-188-13	Sequence 13, Appl
33	37	52.9	16	5	PCT-US94-07019-2	Sequence 2, Appli
34	37	52.9	16	5	PCT-US94-07019-13	Sequence 13, Appl
35	37	52.9	17	2	US-08-569-188-4	Sequence 4, Appli
36	37	52.9	17	2	US-08-569-188-14	Sequence 14, Appl
37	37	52.9	17	5	PCT-US94-07019-4	Sequence 4, Appli
38	37	52.9	17	5	PCT-US94-07019-14	Sequence 14, Appl
39	37	52.9	18	2	US-08-569-188-6	Sequence 6, Appli
40	37	52.9	18	2	US-08-569-188-15	Sequence 15, Appl
41	37	52.9	18	4	US-08-960-054A-12	Sequence 12, Appl
42	37	52.9	18	4	US-08-958-993A-12	Sequence 12, Appl
43	37	52.9	18	4	US-09-296-089-36	Sequence 36, Appl
44	37	52.9	18	5	PCT-US94-07019-6	Sequence 6, Appli
45	37	52.9	18	5	PCT-US94-07019-15	Sequence 15, Appl

ALIGNMENTS

```
RESULT   1
US-09-000-692-1
; Sequence 1, Application US/09000692
; Patent No. 6339067
; GENERAL INFORMATION:
;  APPLICANT: WOLFF, JON A
;  APPLICANT: HAGSTROM, JAMES E
;  APPLICANT: BUDKER, VLADIMIR G
;  APPLICANT: TRUBETSKOY, VLADIMIR S
;  APPLICANT: SLATTUM, PAUL M
;  APPLICANT: HANSON, LISA J
;  TITLE OF INVENTION: A PROCESS OF MAKING A COMPOUND BY FORMING A POLYMER
;  TITLE OF INVENTION: FROM A TEMPLATE  DRUG
;  FILE REFERENCE: TPCIP000692
;  CURRENT APPLICATION NUMBER: US/09/000,692
;  CURRENT FILING DATE: 1997-12-30
;  EARLIER APPLICATION NUMBER: 08/778657
;  EARLIER FILING DATE: 1997-01-03
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism:AMPHIPATHIC
US-09-000-692-1


  Query Match             64.3%;  Score 45;  DB 4;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 1;
  Matches   12;  Conservative    0;  Mismatches    2;  Indels    6;  Gaps    1;

Qy      1 CKLLLKL------LLKLLKC 14
          |||| ||      ||| |||
Db      1 CKLLKKLLKLWKKLLKKLKC 20


RESULT   2
US-08-569-188-8
; Sequence 8, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
```

```
;     STATE:  DELAWARE
;     COUNTRY:  UNITED STATES OF AMERICA
;     ZIP:  19898
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  DISKETTE, 3.50 INCH
;     COMPUTER:  IBM PC COMPATIBLE
;     OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;     SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/569,188
;     FILING DATE:
;     CLASSIFICATION:  525
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/082,852
;     FILING DATE:  JUNE 22, 1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  LINDA AXAMETHY FLOYD
;     REGISTRATION NUMBER:  33,692
;     REFERENCE/DOCKET NUMBER:  CR-9295-A
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  302-892-8112
;     TELEFAX:  302-773-0164
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  peptide
US-08-569-188-8


  Query Match                 57.1%;  Score 40;  DB 2;  Length 14;
  Best Local Similarity       83.3%;  Pred. No. 3.9;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         2 KLLLKLLLKLLK 13
           | ||||| ||||
Db         2 KKLLKLLKKLLK 13


RESULT   3
PCT-US94-07019-8
; Sequence 8, Application PC/TUS9407019
;   GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;     TITLE OF INVENTION:  COMPOSITIONS
;     NUMBER OF SEQUENCES:  15
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  MACINTOSH
;       OPERATING SYSTEM:  MACINTOSH 6.0
;       SOFTWARE:  MICROSOFT WORD, 4.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/07019
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/082,852
;       FILING DATE:  JUNE 22, 1993
;     INFORMATION FOR SEQ ID NO:  8:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH:  14 amino acids
;         TYPE:  amino acid
;         STRANDEDNESS:  unknown
;         TOPOLOGY:  unknown
;       MOLECULE TYPE:  peptide
PCT-US94-07019-8


  Query Match                 57.1%;  Score 40;  DB 5;  Length 14;
  Best Local Similarity       83.3%;  Pred. No. 3.9;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         2 KLLLKLLLKLLK 13
           | ||||| ||||
Db         2 KKLLKLLKKLLK 13


RESULT   4
US-08-569-188-1
; Sequence 1, Application US/08569188
; Patent No. 5847047
;   GENERAL INFORMATION:
;     APPLICANT:  SHARON LPRETTA HAYNIE
;     TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;     NUMBER OF SEQUENCES:  18
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;       STREET:  1007 MARKET STREET
;       CITY:  WILMINGTON
;       STATE:  DELAWARE
;       COUNTRY:  UNITED STATES OF AMERICA
;       ZIP:  19898
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  DISKETTE, 3.50 INCH
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;       SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/569,188
;       FILING DATE:
;       CLASSIFICATION:  525
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/082,852
;       FILING DATE:  JUNE 22, 1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  LINDA AXAMETHY FLOYD
;       REGISTRATION NUMBER:  33,692
;       REFERENCE/DOCKET NUMBER:  CR-9295-A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  302-892-8112
;       TELEFAX:  302-773-0164
;     INFORMATION FOR SEQ ID NO:  1:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH:  16 amino acids
;         TYPE:  amino acid
;         STRANDEDNESS:  unknown
;         TOPOLOGY:  unknown
;       MOLECULE TYPE:  peptide
US-08-569-188-1


  Query Match                 57.1%;  Score 40;  DB 2;  Length 16;
  Best Local Similarity       83.3%;  Pred. No. 4.3;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         2 KLLLKLLLKLLK 13
           | ||||| ||||
Db         4 KKLLKLLKKLLK 15


RESULT   5
US-08-569-188-10
; Sequence 10, Application US/08569188
; Patent No. 5847047
;   GENERAL INFORMATION:
;     APPLICANT:  SHARON LPRETTA HAYNIE
;     TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;     NUMBER OF SEQUENCES:  18
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;       STREET:  1007 MARKET STREET
;       CITY:  WILMINGTON
;       STATE:  DELAWARE
;       COUNTRY:  UNITED STATES OF AMERICA
;       ZIP:  19898
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-10


  Query Match             57.1%;  Score 40;  DB 2;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 4.3;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | |||||  ||||
Db       4 KKLLKLLKKLLK 15


RESULT   6
US-08-569-188-11
; Sequence 11, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-11


  Query Match             57.1%;  Score 40;  DB 2;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 4.3;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | |||||  ||||
Db       4 KKLLKLLKKLLK 15


RESULT   7
US-08-569-188-12
; Sequence 12, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-12


  Query Match             57.1%;  Score 40;  DB 2;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 4.3;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | |||||  ||||
Db       4 KKLLKLLKKLLK 15


RESULT   8
PCT-US94-07019-1
```

```
; Sequence 1, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-1


  Query Match             57.1%;  Score 40;  DB 5;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 4.3;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15


RESULT   9
PCT-US94-07019-10
; Sequence 10, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-10


  Query Match             57.1%;  Score 40;  DB 5;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 4.3;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15


RESULT  10
PCT-US94-07019-11
; Sequence 11, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-11


  Query Match             57.1%;  Score 40;  DB 5;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 4.3;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15


RESULT  11
PCT-US94-07019-12
; Sequence 12, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-12


  Query Match             57.1%;  Score 40;  DB 5;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 4.3;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | ||||| ||||
Db       4 KKLLKLLKKLLK 15
```

```
SULT 12
S-08-569-188-3
; Sequence 3, Application US/08569188
; Patent No. 5847047
;  GENERAL INFORMATION:
;    APPLICANT:  SHARON LPRETTA HAYNIE
;    TITLE OF INVENTION:  NOVEL ANTIMICORBIAL COMPOSITIONS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,188
;      FILING DATE:
;      CLASSIFICATION:  525
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LINDA AXAMETHY FLOYD
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9295-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-569-188-3


  Query Match             57.1%;  Score 40;  DB 2;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 4.6;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | |||||  ||||
Db       5 KKLLKLLKKLLK 16


RESULT 13
US-08-818-253-39
; Sequence 39, Application US/08818253
; Patent No. 5998204
;  GENERAL INFORMATION:
;    APPLICANT:  Tsien, Roger Y.
;    APPLICANT:  Miyawaki, Atsushi
;    TITLE OF INVENTION:  FLUORESCENT PROTEIN SENSORS FOR
;    TITLE OF INVENTION:  DETECTION OF ANALYTES
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,253
;      FILING DATE:  14-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Ph.D., Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07257/043001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-818-253-39


  Query Match             57.1%;  Score 40;  DB 2;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 4.6;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | |||||  ||||
Db       4 KKLLKLLKKLLK 15


RESULT 14
US-08-818-252-39
; Sequence 39, Application US/08818252B
; Patent No. 6197928
; GENERAL INFORMATION:
;  APPLICANT: Tsien, Roger Y.
;  APPLICANT: Miyawaki, Atsushi
;  TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
;  TITLE OF INVENTION: DETECTION OF ANALYTES
;  FILE REFERENCE: 07257/042001
;  CURRENT APPLICATION NUMBER: US/08/818,252B
;  CURRENT FILING DATE: 1997-03-14
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Calmodulin binding peptide-2
US-08-818-252-39


  Query Match             57.1%;  Score 40;  DB 4;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 4.6;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
           | |||||  ||||
Db       4 KKLLKLLKKLLK 15


RESULT 15
PCT-US94-07019-3
; Sequence 3, Application PC/TUS9407019
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  NOVEL ANTIMICROBIAL
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  15
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  MACINTOSH
;      OPERATING SYSTEM:  MACINTOSH 6.0
;      SOFTWARE:  MICROSOFT WORD, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07019
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/082,852
;      FILING DATE:  JUNE 22, 1993
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US94-07019-3


  Query Match             57.1%;  Score 40;  DB 5;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 4.6;
  Matches   10;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 KLLLKLLLKLLK 13
            | ||||| ||||
Db       5 KKLLKLLKKLLK 16


Search completed: June 17, 2002, 12:42:05
Job time: 225 sec
```